DOCUMENT RESUME

ED 062 780 EM 009 777

AUTHOR	Moloney, James Michael
TITLE	An Investigation of College Student Performance on a Logic Curriculum in a Computer-Assisted Instruction Setting.
INSTITUTION	Stanford Univ., Calif. Inst. for Mathematical Studies in Social Science.
SPONS AGENCY	National Science Foundation, Washington, D.C.
REPORT NO	TR-183
PUB DATE	28 Jan 72
NOTE	99p.; Psychology & Education Series
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Complexity Level; *Computer Assisted Instruction; *Logic; Multiple Choice Tests; Problems; *Problem Solving; *Teaching Techniques
IDENTIFIERS	LIS; Logic Instruction System

ABSTRACT

The objective of this study is to identify some of the structural features of an elementary logic curriculum which affect logic problem difficulty. The system under review is a computer-based logic instructional system (LIS) at Stanford University. Four modes of problem presentation--multiple-choice, truth-analysis, counterexample, and derive--are described. Various empirical measures of problem difficulty and measures of problem structure (including structural variables, standard proof variables, and sequential variables) are considered. The performance of college students using the system is analyzed, and variables which contribute to the difficulty of a problem are identified. (JK)

ED 062780

N-A1

AN INVESTIGATION OF COLLEGE STUDENT PERFORMANCE ON A LOGIC CURRICULUM IN A COMPUTER-ASSISTED INSTRUCTION SETTING

BY

JAMES MICHAEL MOLONEY

SCOPE OF INTEREST NOTICE
The ERIC Facility has assigned this document for processing to: EM

In our judgement, this document is also of interest to the clearinghouses noted to the right. Indexing should reflect their special points of view. HE

TECHNICAL REPORT NO. 183

JANUARY 28, 1972

PSYCHOLOGY & EDUCATION SERIES

INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

EM 009 111

TECHNICAL REPORTS

PSYCHOLOGY SERIES

INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES

(Place of publication shown in parentheses; if published title is different from title of Technical Report,
this is also shown in parentheses.)

(For reports no. 1 - 44, see Technical Report no. 125.)

50 R. C. Atkinson and R. C. Calfee. Mathematical learning theory. January 2, 1963. (In B. B. Wolman (Ed.), Scientific Psychology. New York: Basic Books, Inc., 1965. Pp. 254-275)

51 P. Suppes, E. Crothers, and R. Weir. Application of mathematical learning theory and linguistic analysis to vowel phoneme matching in Russian words. December 28, 1962.

52 R. C. Atkinson, R. Calfee, G. Sommer, W. Jeffrey and R. Shoemaker. A test of three models for stimulus compounding with children. January 29, 1963. (J. exp. Psychol., 1964, 67, 52-58)

53 E. Crothers. General Markov models for learning with inter-trial forgetting. April 8, 1963.

54 J. L. Myers and R. C. Atkinson. Choice behavior and reward structure. May 24, 1963. (Journal math. Psychol., 1964, 1, 170-203)

55 R. E. Robinson. A set-theoretical approach to empirical meaningfulness of measurement statements. June 10, 1963.

56 E. Crothers, R. Weir and P. Palmer. The role of transcription in the learning of the orthographic representations of Russian sounds. June 17, 1963.

57 P. Suppes. Problems of optimization in learning a list of simple items. July 22, 1963. (In Maynard W. Shelly, II and Glenn L. Bryan (Eds.), Human Judgments and Optimality. New York: Wiley. 1964. Pp. 116-126)

58 R. C. Atkinson and E. J. Crothers. Theoretical note: all-or-none learning and intertrial forgetting. July 24, 1963.

59 R. C. Calfee. Long-term behavior of rats under probabilistic reinforcement schedules. October 1, 1963.

60 R. C. Atkinson and E. J. Crothers. Tests of acquisition and retention, axioms for paired-associate learning. October 25, 1963. (A comparison of paired-associate learning models having different acquisition and retention axioms, J. math. Psychol., 1964, 1, 285-315)

61 W. J. McGill and J. Gibbon. The general-gamma distribution and reaction times. November 20, 1963. (J. math. Psychol., 1965, 2, 1-18)

62 M. F. Norman. Incremental learning on random trials. December 9, 1963. (J. math. Psychol., 1964, 1, 336-351)

63 P. Suppes. The development of mathematical concepts in children. February 25, 1964. (On the behavioral foundations of mathematical concepts. Monographs of the Society for Research in Child Development, 1965, 30, 60-96)

64 P. Suppes. Mathematical concept formation in children. April 10, 1964. (Amer. Psychologist, 1966, 21, 139-150)

65 R. C. Calfee, R. C. Atkinson, and T. Shelton, Jr. Mathematical models for verbal learning. August 21, 1964. (In N. Wiener and J. P. Schoda (Eds.), Cybernetics of the Nervous System: Progress in Brain Research. Amsterdam, The Netherlands: Elsevier Publishing Co., 1965. Pp. 333-349)

66 L. Keller, M. Cole, C. J. Burke, and W. K. Estes. Paired associate learning with differential rewards. August 20, 1964. (Reward and information values of trial outcomes in paired associate learning. (Psychol. Monogr., 1965, 79, 1-21)

67 M. F. Norman. A probabilistic model for free-responding. December 14, 1964.

68 W. K. Estes and H. A. Taylor. Visual detection in relation to display size and redundancy of critical elements. January 25, 1965, Revised 7-1-65. (Perception and Psychophysics, 1966, 1, 9-16)

69 P. Suppes and J. Donio. Foundations of stimulus-sampling theory for continuous-time processes. February 9, 1965. (J. math. Psychol., 1967, 4, 202-225)

70 R. C. Atkinson and R. A. Kinchla. A learning model for forced-choice detection experiments. February 10, 1965. (Br. J. math stat. Psychol., 1965, 18, 184-206)

71 E. J. Crothers. Presentation orders for items from different categories. March 10, 1965.

72 P. Suppes, G. Groen, and M. Schlag-Rey. Some models for response latency in paired-associates learning. May 5, 1965. (J. math. Psychol., 1966, 3, 99-128)

73 M. V. Levine. The generalization function in the probability learning experiment. June 3, 1965.

74 D. Hansen and T. S. Rodgers. An exploration of psycholinguistic units in initial reading. July 6, 1965.

75 B. C. Arnold. A correlated urn-scheme for a continuum of responses. July 20, 1965.

76 C. Izawa and W. K. Estes. Reinforcement-test sequences in paired-associate learning. August 1, 1965. (Psychol. Reports, 1966, 18, 879-919)

77 S. L. Blehart. Pattern discrimination learning with Rhesus monkeys. September 1, 1965. (Psychol. Reports, 1966, 19, 311-324)

78 J. L. Phillips and R. C. Atkinson. The effects of display size on short-term memory. August 31, 1965.

79 R. C. Atkinson and R. M. Shiffrin. Mathematical models for memory and learning. September 20, 1965.

80 P. Suppes. The psychological foundations of mathematics. October 25, 1965. (Colloques Internationaux du Centre National de la Recherche Scientifique. Editions du Centre National de la Recherche Scientifique. Paris: 1967. Pp. 213-242)

81 P. Suppes. Computer-assisted instruction in the schools: potentialities, problems, prospects. October 29, 1965.

82 R. A. Kinchla, J. Townsend, J. Yellott, Jr., and R. C. Atkinson. Influence of correlated visual cues on auditory signal detection. November 2, 1965. (Perception and Psychophysics, 1966, 1, 67-73)

83 P. Suppes, M. Jerman, and G. Groen. Arithmetic drills and review on a computer-based teletype. November 5, 1965. (Arithmetic Teacher, April 1966, 303-309.

84 P. Suppes and L. Hyman. Concept learning with non-verbal geometrical stimuli. November 15, 1968.

85 P. Holland. A variation on the minimum chi-square test. (J. math. Psychol., 1967, 3, 377-413).

86 P. Suppes. Accelerated program in elementary-school mathematics -- the second year. November 22, 1965. (Psychology in the Schools, 1966, 3, 294-307)

87 P. Lorenzen and F. Binford. Logic as a dialogical game. November 29, 1965.

88 L. Keller, W. J. Thomson, J. R. Tweedy, and R. C. Atkinson. The effects of reinforcement interval on the acquisition of paired-associate responses. December 10, 1965. (J. exp. Psychol., 1967, 73, 268-277)

89 J. I. Yellott, Jr. Some effects on noncontingent success in human probability learning. December 15, 1965.

90 P. Suppes and G. Groen. Some counting models for first-grade performance data on simple addition facts. January 14, 1966. (In J. M. Scandura (Ed.), Research in Mathematics Education. Washington, D. C.: NCTM, 1967. Pp. 35-43.

91 P. Suppes. Information processing and choice behavior. January 31, 1966.

92 G. Groen and R. C. Atkinson. Models for optimizing the learning process. February 11, 1966. (Psychol. Bulletin, 1966, 66, 309-320)

93 R. C. Atkinson and D. Hansen. Computer-assisted instruction in initial reading: Stanford project. March 17, 1966. (Reading Research Quarterly, 1966, 2, 5-25)

94 P. Suppes. Probabilistic inference and the concept of total evidence. March 23, 1966. (In J. Hintikka and P. Suppes (Eds.), Aspects of Inductive Logic. Amsterdam: North-Holland Publishing Co., 1966. Pp. 49-65.

95 P. Suppes. The axiomatic method in high-school mathematics. April 12, 1966. (The Role of Axiomatics and Problem Solving in Mathematics. The Conference Board of the Mathematical Sciences, Washington, D. C. Ginn and Co., 1966. Pp. 69-76.

(Continued on inside back cover)

AN INVESTIGATION OF COLLEGE STUDENT PERFORMANCE ON A LOGIC CURRICULUM IN A COMPUTER-ASSISTED INSTRUCTION SETTING

by

James Michael Moloney

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

TECHNICAL REPORT NO. 183

January 28, 1972

PSYCHOLOGY AND EDUCATION SERIES

Reproduction in Whole or in Part Is Permitted for

Any Purpose of the United States Government

INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

ACKNOWLEDGMENTS

I wish to express my thanks to Dr. Patrick Suppes both for suggesting the problem and for his supervision and encouragement in the research that led to this dissertation. I would also like to thank the other members of my dissertation committee, Drs. Janet Elashoff and Richard Atkinson, for their aid and suggestions.

Finally, I am grateful to Mr. Michael Kane who assisted in the data collection and reduction.

This research was supported by National Science Foundation Grant NSFG-443X.

TABLE OF CONTENTS

Page

Acknowledgments i

Table of Contents ii

Chapter

I. Introduction 1

II. Definition of the Problem 6

III. Description of the Study and of the Models 21

IV. Results 30

V. Discussion 64

References 71

Appendix A. Examples of the Different Problems on the LIS 73

Appendix B. A List of the Rules of Inference, Theorems and Axioms Used in LIS 77

Appendix C. Two Examples of the Derivation Problems from LIS 83

Appendix D. Description of Data Analysis Programs 89

CHAPTER I

INTRODUCTION

The investigation described in this dissertation is partially motivated by a desire to focus attention on certain deficiencies in computer-assisted instruction (CAI) research. The current emphasis in CAI research is on exploring and discovering new ways in which humans and computers can interact. This involves the design of special hardware and the implementation of new programming techniques (software). The reader is referred to Wexler (1970) for a brief account of the historical development of CAI. In his account, one can clearly see that the primary emphasis in CAI research projects has been system development. Most of the projects have been implemented under ideal operating conditions for small and highly motivated groups of students, while little or no attention has been given to evaluating the curriculum or the pedagogical methods used.

Usually, the system is designed to simulate some intuitive concept of a "good teacher" and to "individualize instruction." The result has been a large collection of complex, interesting, and, from a computer scientist's point of view, valuable instructional systems. However, little machinery is available to judge their educational value or relevance in any systematic or quantitative way.

For several years, the CAI Laboratory at the Institute for Mathematical Studies in the Social Sciences (IMSSS) has been offering a course in mathematical logic. The availability of this course has made it possible to collect information on student behavior in elementary mathematical logic, information which was unavailable before the advent of CAI. This dissertation is a first attempt at an in-depth analysis of some of the factors which contribute to problem difficulty in elementary mathematical logic. The major focus of the study is to develop an understanding which might eventually lead to a quantitative theory of problem solving in logic. This work is in the spirit of the analyses in elementary mathematics to be found in Suppes, Jerman and Brian (1968) and in Loftus (1970). My approach involves formally describing the relationship between structural features of logic problems and problem difficulty, as well as the development of models which predict difficulty as a function of curriculum structure.

Unfortunately, the researcher interested in utilizing an operational CAI system faces many novel problems. In the remainder of this chapter I shall mention some of these problems because I consider them important and relevant to discussions of CAI. However, it is not the purpose of this dissertation to provide detailed discussion or present serious evidence on these problems.

Not only does the computer provide us with the ability to create a large number of new educational environments, but it also provides us with a capability for recording and preserving many aspects of student behavior. However, the utilization of this data-collection capability presents several problems. In a large-scale CAI system, such as the one at IMSSS where many CAI programs are being run concurrently, it is possible to become inundated with student-response data. As the volume of data collected increases, system reliability goes down and computer-response time goes up. Also, the time and overhead required to remove the data from high-cost, short-term devices to low-cost, long-term devices increases. It is a serious mistake to overload the system by indiscriminately recording every student response. To minimize the amount of data collected, one must plan carefully what is required for a particular study.

This planning must involve another feature of computer systems, namely, the finite probability of system failure and its effect on the data. Systems can and do fail, and data are unavoidably lost because it is not economically feasible to have duplicate backup facilities for educational systems. As a result, it is not always possible to implement carefully controlled experimental designs or paradigms on a large-scale operational CAI system.

The problems facing the data collector are categorized under three major headings. First, are problems which arise as a result of hardware and/or software failure. The failure of any component may result in a serious curtailment or cessation of operations. Failures usually have an adverse effect on data collection, the chief effect being an unrecoverable loss of a part of the data. Precautions can be taken to minimize the loss of data, but the loss cannot be predicted or entirely prevented.

In a CAI classroom, the second major area of concern is the student-proctor interaction. A proctor is the person who supervises and aids students while they are at the computer terminals. In the CAI system at Stanford University, personnel who serve as proctors vary widely in training and background. In some elementary schools, there are full-time proctors on duty, while at other schools with fewer terminals, the classroom teachers serve as proctors. In college courses, the teaching assistant usually serves as the proctor, or, in some cases, the proctor is a subject-area expert. Since there has been no attempt to set up general guidelines for CAI proctors, they are likely to see their roles differently and, thus, to differ in the amount and kind of aid that they give to their students. If research done in CAI is to have widespread application, then more thought must be given to standardize the proctor's role. If the proctor also happens to be the teacher of the course, he may sometimes see attempts at standardizing procedures (at the terminals) as conflicting with his teaching goals. This conflict must be reduced if we wish to use CAI as a research tool.

The third and final area of concern is curriculum writing, particularly those parts of the curriculum written specifically for the computer. Ideally, from the point of view of the educational researcher, a curriculum should be designed to provide evidence for evaluating the hypotheses on which it is based. Frequently, teachers of the courses and administrators may not share the researcher's zeal for a neat experimental design. Often the curriculum already exists, and curriculum writers are not inclined to rewrite their material for the researcher's sake. In most cases, there is no empirical evidence to convince a teacher or curriculum writer that the changes will be of benefit to his students. Thus, researchers often find it necessary to develop techniques for examining already existing curricula. As the understanding of a particular curriculum grows, the researcher may be able to present more objective reasons why a particular curriculum should be changed. Thus, a curriculum can be changed in ways which are beneficial to the student and to the educational researcher.

Many of the difficulties mentioned above are sufficiently complex to provide, in themselves, the basis for a major study. Therefore, as has been previously state, the scope of this investigation was limited to the extent discussed in Chapter II, namely, the area of curriculum study. I do feel that the problems mentioned in this section are important, and I hope the discussion will stimulate further in-depth study of them.

CHAPTER II

DEFINITION OF THE PROBLEM

The primary objective of this dissertation is to identify some of the structural features of an elementary logic curriculum which affect logic problem difficulty. A related task is to provide an adequate behavioral measure of problem difficulty as well as an objective, quantitative characterization of the curriculum structure. In this chapter, a detailed description of the curriculum under consideration is presented, followed by a discussion of problem difficulty and curriculum structure.

The study involves several aspects of the existing computer-based logic instructional system (LIS) at Stanford University. The term 'logic instructional system' is used to emphasize that this is the investigation of a specific curriculum in the context of a large-scale CAI system. The computer configuration under consideration is a modified Digital Equipment Corporation (DEC) PDP-10 time-sharing system located at IMSSS.

Becoming operational at Stanford in 1963, LIS was originally designed as a self-contained tutorial program to teach sentential logic to bright elementary-school children. It was first implemented on the DEC PDP-1 system at IMSSS, and students traveled from the surrounding elementary-school districts to the instructional laboratory at Stanford to take their logic lessons. Later, the students were able to take their lessons on teletypes located in their schools.

Since its inception, both the curriculum and the program have been under constant modification and revision as new operational modes have been added. In the fall of 1969, a version of LIS was implemented on the PDP-10 system. In the spring of 1970, data-collection routines were added by the author to the PDP-10 version. LIS, as described below, is the current version of the program with data-collection capabilities.

The description of LIS will proceed as follows. First, a brief description of the modes of problem presentation is given. These are multiple-choice, truth-analysis, counterexample, and derive modes. An example of each mode can be found in Appendix A. Next there is a detailed discussion of the types of input which the students are allowed to make and the manner in which LIS handles invalid student input. This is followed by a discussion of the program clocks. Finally, an outline of the subject matter of the LIS curricula is presented. Since this study is concerned primarily with student performance, it is not appropriate to include a detailed description of the organization and logic of the operating program.

The multiple-choice mode needs little explanation. Students are presented with a small body of text. The text is usually an explanation of a concept followed by a question, or else it is a question on some previously explained material. Then two or three lettered responses are presented, and the student is required to type in the letter corresponding to the correct response. If he types in the correct response, the computer types correct and presents the next problem. If he types an incorrect response, the computer types wrong, try again. This continues until the student enters a correct response.

In the truth-analysis mode, the student is required to compute the truth value of a formula. In one form of truth analysis, the machine assigns the value T or the value F to each sentential variable occurring in the formula and then presents the student with each subformula. The student types the truth value for each of these subformulas. After he has assigned values to all subformulas, he is presented with the whole formula, and he must type in its truth value. If his answer is correct, he receives the next problem. If it is not, he must repeat the problem.

In the other form of truth-analysis problem the student is given the truth value of the conclusions. His task is to assign truth values to the sentence letters such that the conclusion takes on its given value. As in the other type, the problem is repeated until the student makes the correct truth assignments.

The counterexample mode is similar to the truth-analysis mode. The student is presented with a formula and zero or more premises and asked to make truth assignments such that the premises are true and the conclusion false. He is presented with each variable, and he assigns a truth value to it. Using his assignments, he computes the truth values of each subformula and then of each premise. If any premise is found to be false, he is required to restart the problem. If the premises are true, he is presented with the conclusion and asked to compute its truth value. If the conclusion is false, the computer types correct, and he is presented with the next problem. If the conclusion is found to be true, he must restart the problem.

In the derive mode, the student is required to construct a derivation. For this purpose, he has at his disposal a large number of rules of inference, axioms, and, eventually, theorems. A list of these rules can be found in Appendix B. (For the sake of brevity, we use the term 'rule' to denote 'rule of inference,' 'axiom,' and 'theorem' for the remainder of this dissertation.) The student is permitted to type any rule which is logically valid at any step in a derivation. The rule need not, in any sense, bring the student closer to the desired conclusion. Thus, as long as the student continues to enter logically valid rules, he is free to use any line of reasoning that he wishes. At present, there is a 32-line limit on the length of a derivation, but for the problems considered here this restriction is inconsequential.

Except for the rules IP, FIN, and DLL, each rule has the form $n_1.n_2X_1X_2n_3$, where n_1, n_2, and n_3 are either integers or null, X_1 is a letter of the alphabet, and X_2 is a letter or null. To provide an illustration of the way in which rules are used, we have included Appendix C. It contains two typical derivation problems. The first example is from sentential logic and contains an instance of the rule IP. The second example is a typical algebra problem.

The rule DLL (delete last line) allows the student to "erase" his last line. When the student types DLL, the computer deletes all of its internal references to the line previously entered by the student. The next line entered by the student is given the same number as the last line deleted. The student is permitted to delete, sequentially, any line that he has entered.

If a student attempts to enter a rule which is not logically valid or to enter a nonexistent rule or an improper rule format, he is given an error message. These are one- or two-line messages typed to the student which explain the nature of his mistake. Some typical error messages are included in Example 2 of Appendix C.

A fifth type of problem presentation asks the student to find either a derivation or a counterexample (problem 505.25, Appendix A). The student must decide whether the formula presented is true or false. If he decides that a counterexample exists, he type CEX and the machine enters the counterexample mode; otherwise, he types DER, and the machine enters the derive mode. In either case, the computer does not evaluate his choice. That is, if he types CEX and a counterexample does not exist, he is still permitted to try to find one, and vice versa.

There are three clocks in LIS which are relevant to this discussion. These clocks may be thought of as alarm clocks. They are set by the program to "ring" or "fire" after some specific duration. When a clock fires, it signals the program to initiate some particular action.

Some problems contain hints which are stored with the problem in the problem file. If a student desires help, he may type H. A hint is available only if one has been written for the problem and the hint clock has fired. If a hint exists for a problem, but the clock has not fired when a student types H, he is told to wait a little longer. If there is no hint for a problem and the student asks for help, he is told that no hint is available. The hint clock is set to fire 0.5 minutes after the beginning of a problem and after each response.

The problem clock is set to fire two minutes after the last student input. If the student inputs any character prior to this time, the problem clock is reset. If the clock fires, the student is automatically signed off the terminal and his session is terminated.

The session clock is set when the student signs on. It fires fifty minutes later. The student is then signed off at the completion of his current problem, although the student may sign himself off, at any time, by typing FIN. He is, of course, free to sign back on again at any time, and then his session clock is reset to fifty minutes.

The logic curriculum is arranged by lesson. Each lesson contains a different number of problems and is designed to teach one or more concepts. There are five series of lessons. The 100 and 200 series lessons were designed for elementary and junior high school students. The 400, 500 and 600 series lessons were designed primarily for college students.

The 400 and 500 series lessons concentrate on the axioms for an ordered field. The student begins with a review of sentential logic. He is then given a set of axioms for addition of numbers that includes commutativity, associativity, and the properties of zero and negative numbers. Using the axioms and rules of inference, he derives a number of theorems on the addition of numbers. After a theorem has been proved by a student, it becomes available to him for use in later proofs. Following the section on addition, a similar treatment is given to multiplication and fractions. The student next studies some properties of the ordering relation "less than." The final section gives the same axiomatic treatment to the Boolean or class algebra.

The 500 series concentrates on the review of sentential logic. This series was implemented primarily to give the student practice in presenting counterexamples to unsound arguments. It is the only series of the college curriculum in which counterexample-mode problems can be found.

Finally, the 600 series was added in the fall of 1970 for use in Philosophy 3, The Logic of Political Argument. It was designed to adhere in structure as closely as possible to the 500 series. Only the semantic content of some of the problems was changed.

This study is concerned with examining the relationship between the structural properties of logic problems and problem difficulty, expressed as a function of student performance. In earlier studies of this nature on elementary mathematics curricula, the proportion of students who successfully completed a problem was used as a measure of difficulty. In these earlier studies, the problems were such that the correctness of a single response was a good indication of whether the student had successfully performed the task required. In logic, the "correct answer" or the derived expression is not the object of interest. The student must present evidence that he has constructed a valid argument. The evidence takes the form of a valid derivation using the rules of LIS. Further, the student is not permitted to advance to the next problem until he has successfully completed his current one. Thus, it would not be useful or meaningful to use proportion correct as an indicator of problem difficulty. I had to look for other, less obvious, measures of problem difficulty.

In the search for a measure of difficulty I was constrained to quantities measurable by our system. Since this was an investigation of a college curriculum under actual teaching conditions, it was desirable to make the data collection invisable to the student. Thus, the data available were the characters which the student typed to the system, the times at which these characters were entered and the system's response to the student. In the ensuing paragraphs I consider some of the alternative measures of difficulty, definable in terms of the information at our disposal.

First, the mean number of error messages per problem can be used as a measure of problem difficulty. However, there are several possible explanations why a student may enter a response which generates an error message. First, error messages may occur as a result of typing errors, such as a student accidentally hitting the wrong key. Second, a student may know which rule he needs to proceed but he may be unsure of how to enter it in LIS. Or third, he may, in fact, have a faulty understanding of a rule. To gain a more complete understanding of the reasons behind behavior which results in error messages would require a far deeper analysis of error messages than is planned for this study. It is also relevant to note that a student may be unable to do a problem and yet generate no error messages. He can do this either by having no input at all or by inputting rules which he knows, but which are irrelevant to a correct derivation. However, the relationship of this measure with the other measures defined below was examined. This measure will be referred to as variable B5, in order to remain consistent with order in which the variables were listed by the data reduction programs.

Next, consider the number of lines in the derivation--that is, the number of correctly entered rules for a valid derivation. The measure of difficulty is defined as the mean number of lines per proof per problem and referred to as varible Bl. This criterion of difficulty has two serious drawbacks. First, a proof for a problem may be very short, yet the problem is considered, intuitively, difficult. Problems which require "tricks" or unusual approaches fall into this category. Second, problems which require a large number of lines are sometimes considered intuitively easy. These are problems which require straightforward applications of familiar rules.

Third, consider the elapsed time from the start of a problem to its solution. Define as a measure of difficulty the mean latency to completion and denote it by B2. More precisely, the latency is the sum of the latencies for each valid line entered by the student. (See Appendix D for a more detailed description.) Unfortunately, one of the objections stated in the previous paragraph may be applied to this measure also. Latency is an increasing function of the number of lines in a proof. Thus, "easy" problems which require many lines will have large latencies. As a result, I was not able to distinguish between short, "tricky" problems and longer, straightforward ones.

It seems more reasonable to believe that problem difficulty is some function of problem length and latency. Thus, a fourth possibility is the mean latency per line. This quantity is defined in two ways. Variable B3 is defined as

$$B3 = \sum_{i=1}^{N} \left(\frac{T_i}{L1_i} \right) \Big/ N ,$$

where $L1_i$ is the number of valid lines entered by student i, T_i is total latency to solution for student i and N is the number of students solving the problem. Variable B4 is defined as

$$B4 = \sum_{i=1}^{N} \left(\frac{T_i}{L2_i} \right) \Big/ N ,$$

where $L2_i$ is $(L1 - 2*DLL_i)$, DLL_i is the number of occurrences of the rule DLL in the proof of student i and T_i and N are as above. Both of these measures are free from the objections mentioned above and agree with one's intuitive feelings of problem difficulty.

It is shown in Chapter IV that variables B3 and B4 are highly correlated (.99). Note that B3 includes the false starts or irrelevant paths which the student has decided to "erase" from his proof by the use of DLL. B4, on the other hand, includes only those lines which the student has decided will comprise his actual proof. Because of the close relationship between these variables, variable B4 was chosen the measure of difficulty in order to decrease problems of interpretation in the analysis. Thus, the measure of difficulty is the corrected mean latency per line.

Having defined the emperical measures of problem difficulty, we now turn to a discussion of the variables or structural features of the problems, which are indicators of problem difficulty. These variables must be defined solely in terms of problem and/or curriculum structure and not as a function of the student's performance. The variables are divided into three distinct categories: (a) structural variables, (b) "standard proof" variables, and (c) sequential variables. Each category is discussed separately.

Structural variables are those features of a problem which can be identified by visually examining the problem. These variables are defined solely in terms of the symbols which appear on the teletype prior to student input. A brief description of each follows.

1. The number of words in the problem. This is essentially a measure of the amount of information to be processed by the student. Symbolic logical connectives (V,&,¬,→), arithmetic operators, sentence letters, algebraic variables, numerals, and parentheses are considered as one word each. In studies on elementary-school

mathematics (Suppes, Jerman and Brian, 1968), an analogous variable was significant in predicting performance on problems.

2. The number of symbols in the sentence to be derived. This variable is intended to give one measure of the logical complexity of the problem. The procedure for obtaining a value for (2) is illustrated by the following example. Suppose the problem is

   DERIVE: $A < (5+4)+1 \rightarrow A < 5+((1+3)+1)$.

   There are 23 symbols in the sentence, thus the value of the variable is 23.

3. Number of occurrences of logical connectives in the sentence to be derived. This variable is a slightly different measure of the logical complexity of the problem. To illustrate the procedure for obtaining the value of (3), consider the following simple example:

   DERIVE: $(R\&S) \rightarrow R$.

   There are two logical connectives, namely, & and →. Thus, the value of the variable is 2.

4. The depth of nesting of the most deeply parenthesized expression in the sentence to be derived. This variable is intended to reflect another aspect of logical complexity. The value of this variable is found by counting the number of left parentheses in each expression of the sentence to be derived and choosing the maximum value. If there are no

parentheses, the value is zero. To illustrate this variable, consider again the problem:

DERIVE: $A < (5+4)+1 \rightarrow A < 5+((1+3)+1)$.

We find three parenthesized expressions, namely, (5+4), ((1+3)+1) and (1+3). (5+4) has one left parenthesis, ((1+3)+1) has two left parentheses and (1+3) has one left parenthesis. The maximum value is two, thus the value of the variable is 2.

5. The number of premises. This variable gives some measure of the amount of information a student must take into account and use while attempting a derivation. It seems reasonable to assume that, as the number or premises increases, difficulty will also increase.
6. Problem context (0,1). This variable is a reflection of the context in which the problem occurs. The variable has value one if it is a 500 series problem and zero otherwise.
7. Explanatory material and/or a hint in the problem statement (0,1). The variable has value one if the problem contains explanatory material, zero otherwise.

The "standard proof" variables have an element of subjectivity in their definitions which the first group does not have. They require the availability of a solution or proof for the problem. Since the solution to a logic problem is not unique, there will be some degree of arbitrariness in the selection of a "standard proof." For purposes of this study, those proofs generated by the author will be considered standard.

Several criteria were used by the author in generating the standard proofs. First, the author worked through the entire set of problems included in this study two times. The proofs generated the second time through are used as standard. An attempt was made to construct proofs with a minimal number of lines. Also, within the constraint of producing a minimal proof, an attempt was made to use rules and theorems most recently introduced, wherever possible. It is the judgment of the author that the great majority of the proofs produced are minimal in the sense of containing the least possible number of lines.

It is true that from a mathematical standpoint, it might be desirable to demonstrate that the proofs are minimal. However, the proofs are surely minimal in the majority of cases given and explicit proof would make very little change in the interpretation of my results.

All but one of the "standard proof" variables are the number of occurrences of certain rules used in the standard proof. These rules are:

8. Affirm the antecedent. (AA)
9. Conditional proof. (CP)
10. Indirect proof. (IP)
11. Any axiom.
12. Any theorem. The material included in this study contained only Theorems 1 through 6.
13. The number of lines in the proof.

The third group of variables is made up of the sequential variables. These variables are meant to measure the effect of position of the problem in the curriculum. It is reasoned that the greater the number of rules available to the student, the more difficulty he will have in deciding which one to use. Also, the number of problems completed will affect performance. The following variables are an attempt to quantify these facts.

The first three are simply the number of rules, theorems, and axioms available to the student for the problem. That is, the magnitude of the number of available:

14. Rules of inference.
15. Theorems.
16. Axioms.

The next and final variable provides a measure of the "learning" for each rule. It is defined as:

17. The number of problems since the last introduction of a rule. This variable gives some measure of the amount of practice a student has had on a particular rule.

Table 1 lists the measures discussed in this section. Also included in Table 1 is a transformed variable, denoted S18, which is of importance in the analysis which follows. I have included it in Table 1 in order to provide the reader with a complete list of structural variables used. The significance of variable S18 is discussed in Chapters III and IV.

Insert Table 1 about here

TABLE 1

Behavioral and Structural Variables

I. Measures of Problem Difficulty

- B1. Mean number of lines per derivation
- B2. Mean latency to a correct solution
- B3. Mean latency per line
- B4. Correlated mean latency per line (difficulty)
- B5. Mean number of error messages per derivation

II. Measures of Problem Structure

A. Structural Variables

- S1. Number of words per problem
- S2. Number of symbols in sentence to be derived
- S3. Number of occurences of logical connectives in the sentence to be derived
- S4. Depth of nesting of the most deeply parenthesized expression in the sentence to be derived
- S5. Number of premises
- S6. Problem context
- S7. Inclusion of explanatory material and/or hint of the problem statement

B. Standard Proof Variables

- S8. Number of occurrences of affirm the antecedent (AA)
- S9. Number of occurrences of conditional proof (CP)
- S10. Number of occurrences of indirect proof (IP)
- S11. Number of occurrences of any axiom
- S12. Number of occurrences of any theorem
- S13. Number of lines in the proof

C. Sequential Variables

- S14. Number of rules of inference available
- S15. Number of theorems available
- S16. Number of axioms available
- S17. Number of problems since the last introduction of a rule

D. Transformed Structural Variable

- S18. S5 cubed

CHAPTER III

DESCRIPTION OF THE STUDY AND THE MODELS

In this chapter I discuss the population utilized, outline the method of data collection, describe certain characteristics of the collected data, and outline the methods of analysis. The primary objective of this analysis is to describe in a precise way the relationship between the structural and behavior measures of difficulty and to develop models which will enable us to predict student performance from the structural features of the problems. A secondary objective is to provide some general descriptive information about student performance on the LIS.

The population used in this study consisted of the 27 Stanford University students who enrolled in Philosophy 157 in the summer quarter of 1970, the period during which the data were collected. No special procedures, other than normal departmental prerequisites, were used in the selection of these students. The group consisted only of students who had decided to take the course.

The curriculum under investigation consisted of 203 problems from the computer-based segment of the course. Although the number of students involved in the study is not large, a considerable quantity of information has been collected for each student. Thus, I feel that an ample amount of information is available to successfully accomplish the objectives of this study, even though its generalizability to all student populations is limited.

The students were proctered during their sessions at the terminals by the philosophy graduate students who gave the lecture portion of the course. They received three hours of traditional classroom instruction per week, in addition to the time which they spent at the computer terminal. Also, there was always someone available who was familiar with the computer system and the logic program and who was able to deal with any operational difficulties.

The logic data collection routines were added to the LIS in the spring and early summer of 1970. They were designed and programmed by the author. When the logic program was converted from the PDP-1 to the PDP-10, no provisions for data collection were made. Thus, it was necessary to modify certain sections of an already existing program.

These modifications required several steps. A special data collection routine had to be written in assembly language and interfaced with the logic program. It was decided to store the raw data on disk files during the day and then to transfer each day's data to magnetic tape, where it was kept for later reduction and analysis. The necessary programs were written and debugged in the spring of 1970.

During the time that the data were being collected, some data were lost. As a result of long-term experience with the system (two years), I feel justified in stating that data loss was in no way systematic. However, to support this opinion rigorously would require a much more definitive analysis of the system than is presently available, and I feel that it would be neither feasible nor appropriate to include a detailed analysis of the system in this study.

During the summer of 1970, while the data were being collected, a second series of programs were written by the author. They were designed to convert the raw data into a form acceptable by the standard Biomedical Programs (BMD) used in he final stage of the analysis (Dixon, 1970). These intermediate programs are described in detail in Appendix D.

The results presented in this study were obtained by means of two BMD programs. First, the overall means, standard deviations and correlations of all of the variables described in Chapter II were computed. For this purpose, I modified the BMD06M Canonical Analysis Program to run on the IMSSS PDP-10 system (see Appendix D). An outline of the computational procedure used may be found in the BMD Manual, pp. 207-213. These results are discussed in Chapter IV.

The next step in the analysis was to describe formally the nature and degree of the relationship between the behavioral and structural variables. To do this, the canonical correlations and canonical coefficients were computed by means of BMD06M. Although canonical analysis is a well-known procedure, an outline of the model is provided to avoid any ambiguity in terminology. The development follows that of Morrison (1967).

Consider the two sets of variates: the behavioral variables and the structural variables. Assume that the first set has p variates and the second set has q variates. Suppose that the p + q variates are from some multidimensional population which has been partitioned such that:

$$\underset{\sim}{\mu}' = (\underset{\sim}{\mu}_1, \underset{\sim}{\mu}_2) \qquad \underset{\sim}{\Sigma} = \begin{pmatrix} \underset{\sim}{\Sigma}_{11} & \underset{\sim}{\Sigma}_{12} \\ \underset{\sim}{\Sigma}_{12}' & \underset{\sim}{\Sigma}_{22} \end{pmatrix} .$$

It is assumed that:

1. The elements of $\underset{\sim}{\Sigma}$ are finite.

2. $\underset{\sim}{\Sigma}$ is of full-rank $p + q$.

3. The first $r \leq \min(p,q)$ characteristic roots of

$$\underset{\sim}{\Sigma}^{-1}_{11}\,\underset{\sim}{\Sigma}_{12}\,\underset{\sim}{\Sigma}^{-1}_{22}\,\underset{\sim}{\Sigma}'_{12}$$

are distinct.

From this population, N-observation vectors have been randomly drawn and the sample has been partitioned such that:

$$\underset{\sim}{X}' = (\underset{\sim}{X}_1\,,\ \underset{\sim}{X}_2) \qquad \underset{\sim}{S} = \begin{pmatrix} \underset{\sim}{S}_{11} & \underset{\sim}{S}_{12} \\ \underset{\sim}{S}'_{12} & \underset{\sim}{S}_{22} \end{pmatrix}\ ,$$

in conformance with the above.

We wish to determine the linear compounds

$$\begin{array}{ll} u_1 = \underset{\sim}{a}'_1\,\underset{\sim}{X}_1 & v_1 = \underset{\sim}{b}'_1\,\underset{\sim}{X}_2 \\ u_2 = \underset{\sim}{a}'_2\,\underset{\sim}{X}_1 & v_2 = \underset{\sim}{b}'\,\underset{\sim}{X}_2 \\ \ \ \vdots & \ \ \vdots \\ u_5 = \underset{\sim}{a}'_5\,\underset{\sim}{X}_1 & v_5 = \underset{\sim}{b}'_5\,\underset{\sim}{X}_2\ , \end{array}$$

such that the sample correlation of u_1 and v_1 is greatest, the sample correlation of u_2 and v_2 is greatest among all linear compounds uncorrelated with u_1 and v_1, and so on for all $s = \min(p,q)$ possible pairs.

To do this, solve for λ in

$$|\underset{\sim}{S}_{12}\underset{\sim}{S}^{-1}_{22}\underset{\sim}{S}'_{12} - \lambda\underset{\sim}{S}_{11}| = 0\ .$$

Order the roots from largest to smallest $C_1, C_2, \ldots C_s$. These are the canonical correlations. The coefficients are obtained from the equations

$$(\underset{\sim}{S}_{12}\underset{\sim}{S}^{-1}_{22}\underset{\sim}{S}'_{12} - \underset{\sim}{C}_i\underset{\sim}{S}_{11})\underset{\sim}{a}_i = \underset{\sim}{0}$$

$$(\underset{\sim}{S}'_{12}\underset{\sim}{S}^{-1}_{22}\underset{\sim}{S}_{12} - \underset{\sim}{C}_i\underset{\sim}{S}_{22})\underset{\sim}{b}_i = \underset{\sim}{0}\ ,$$

where a_i and b_i are chosen to satisfy

$$\underset{\sim}{a}'\underset{\sim}{S}_{11}\underset{\sim}{a} = 1 \quad \text{and} \quad \underset{\sim}{b}'\underset{\sim}{S}_{22}\underset{\sim}{b} = 1 \quad .$$

The final stage of the analysis was to fit a regression model in order to predict problem difficulty as a function of problem structure. Although the primary goal was to predict problem difficulty, regression models using B1 and B2 as the dependent variables were also considered. Thus, some idea of the predictive power of the structural variables with respect to these other behavior measures was obtained.

The program used for the regression analysis was the BMD02R. This program was fully implemented on the PDP-10 by IMSSS staff in June, 1970 and further modified by the author (see BMD Manual and Appendix D). Since regression analyses are also a standard statistical procedure, it does not seem appropriate to give a full description of the theory of regression analysis here. However, the model is presented for purposes of developing notation.

The general multiple linear regression model can be written as:

$$\underset{(n \underset{\sim}{x} 1)}{Y} = \underset{(n \underset{\sim}{x} p)}{X} \; \underset{(p \underset{\sim}{x} 1)}{\beta} + \underset{(n \underset{\sim}{x} 1)}{\epsilon}$$

where

$$\underset{\sim}{Y} = \begin{pmatrix} Y_i \\ \vdots \\ Y_n \end{pmatrix} \quad \underset{\sim}{X} = \begin{pmatrix} X_{10} & \cdots & X_{1,\, p-1} \\ \vdots & & \\ X_{np} & & X_{n,\, p-1} \end{pmatrix} \quad \underset{\sim}{\beta} = \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_{p-1} \end{pmatrix} \quad \underset{\sim}{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix} .$$

Y_i are the mean problem difficulties for i-th problem.

X_{ij} are the values of the p-1 structural variables $1 \leq j \leq p-1$.

β_i are the parameters to be estimated.

ϵ_i are the errors.

$X_{i0} = 1$ for all i.

We can write the normal equations as:

$$\underset{\sim}{X}'\underset{\sim}{X}\,\underset{\sim}{\beta} = \underset{\sim}{X}'\underset{\sim}{Y} \ .$$

Assume $E(\underset{\sim}{\varepsilon}) = \underset{\sim}{0}$ and $V(\underset{\sim}{\varepsilon}) = \underset{\sim}{I}\,\sigma^2$, then the least squares estimators $\underset{\sim}{B}$ of $\underset{\sim}{\beta}$ are

$$\underset{\sim}{B} = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\ \underset{\sim}{X}'\underset{\sim}{Y} \ .$$

Assume further that $\underset{\sim}{\varepsilon} \sim N(\underset{\sim}{0}, \underset{\sim}{I}\,\sigma^2)$,

then

$$\underset{\sim}{B} \sim M.V.N.(\underset{\sim}{\beta}, \sigma^2(\underset{\sim}{X}'\underset{\sim}{X})^{-1}).$$

The ANOVA table is shown in Table 2.

Insert Table 2 here

If the model is correct, $MS_{RES} = S^2$ is an estimate of σ^2 . Define the coefficient of multiple determination R^2 as:

$$R^2 = (\underset{\sim}{B}\underset{\sim}{X}'\underset{\sim}{Y} - N\bar{Y}^2)/(\underset{\sim}{Y}'\underset{\sim}{Y} - N\bar{Y}^2) \ .$$

This is usually interpreted as the proportion of variance accounted for by the regression.

I shall now discuss the assumptions made in the regression analysis and the procedures used to check the validity of these assumptions for our data. First, it is assumed that the model is linear in the parameters. Since this was the first attempt at analysis of college student performance on the LIS, no information was available to use as a guide in the selection of a nonlinear model. Thus, until definite information about the form of the relationship between the variables is available, a linear model is assumed.

TABLE 2

Analysis of Variance for Stepwise Multiple Linear Regression

Source	df	SS	MS
Regression	p-1	$\utilde{b}'\utilde{X}'\utilde{Y} - n\bar{Y}^2$	$(\utilde{b}'\utilde{X}'\utilde{Y} - n\bar{Y}^2)/p-1$
Residual	n-p	$\utilde{Y}'\utilde{Y} - \utilde{b}'\utilde{X}'\utilde{Y}$	$(\utilde{Y}'\utilde{Y} - \utilde{b}'\utilde{X}'\utilde{Y})/n-p$
Total	n-1	$\utilde{Y}'\utilde{Y} - nY^2$	

The other assumptions concern the distribution properties of the errors. If $\underset{\sim}{\epsilon}$ is the vector of errors, then assume that $E(\underset{\sim}{\epsilon}) = \underset{\sim}{0}$ and $V(\underset{\sim}{\epsilon}) = \underset{\sim}{I}\,\sigma^2$, that is, the errors are uncorrelated and have common variance. An assumption that the $\underset{\sim}{\epsilon}$ are normally distributed is not necessary to obtain estimates of the parameters, but it is necessary only in order to make tests of statistical significance. These assumptions can be examined by plotting the residuals. The residuals are defined as:

$$R_i = Y_i - \hat{Y}_i \quad .$$

If the fitted model were correct, the residuals should have exhibited tendencies that would seem to confirm the assumptions. My version of the BMD02R program allowed, as optional output, plots of (a) residuals versus computed, (b) residuals versus the independent variables, and (c) dependent variable versus the independent variables. I made all plots in order to determine if the assumptions appeared to be violated. Where the assumptions appeared to be violated, the plots were used to pinpoint the sources of trouble and transformations on the existing variables were used to correct for the violations.

In some cases the assumption $V(\underset{\sim}{\epsilon}) = I\,\sigma^2$ appeared to be violated, perhaps due to the fact that the measure of difficulty is essentially a latency. Again we attempted to remedy this situation by a transformation. Kruskal (1968) discussed a number of variance-stabilizing transformations. The various transformations suggested by Kruskal were considered, and I selected the square-root transformation as the one most useful for my purposes. Kruskal also stated that many authors have remarked that frequently (although not invariably) a single transformation also improves normality, as well as stabilizing variance.

In summary, the analysis was carried out as follows: first, the behavioral data were reduced to a form usable by the standard statistical routines. In the process, we output descriptive summaries of college student performance on the LIS. Next the BMDO6M Canonical Analysis program was used to obtain a concise measure of the relationship between the two sets of variables listed in Figure 1, Chapter II. Finally, using the intuitively best measure of difficulty--correclated mean latency per line--as the dependent variable, I did a stepwise multiple linear regression in order to develop a model which could account for problem difficulty as a function of problem structure.

CHAPTER IV

RESULTS

In this chapter I discuss the results of the analyses which were described in detail in Chapter III. First, the summary statistics for all the variables studied are presented. Next, we discuss some interesting aspects of the data which do not appear in the summary tables. These include a brief discussion of the problems which had extreme values on the behavioral variables. Then we look at the correlations among the variables and discuss the canonical analysis. This section concludes with a discussion of the regression analyses.

Table 3 contains the mean, standard deviation and range for each of the behavioral variables. Table 4 contains these statistics for the structural variables. A brief discussion of several of the values found in the tables will be informative.

Insert Tables 3 and 4 about here

First, in Table 3 note that the means of variables B3 and B4 differ by less than one unit and their ranges are identical. Thus, I have assumed that these variables are slightly different measures for the same underlying behavior. I have chosen to use variable B4 as the "measure of difficulty" for the reasons given in Chapter II (p. 15).
A second interesting aspect of the results is the low error message rate, variable B5. In fact, there were 26 problems for which there was no error at all. This implies that the students were adept at using the rules they had learned.

TABLE 3

Means, Standard Deviations and Ranges for Behavioral Variables

Variable	Mean	Standard Deviation	Low	High
B1	4.55	3.32	1.00	15.80
B2	84.18	85.53	3.97	415.71
B3	15.77	7.54	3.70	47.25
B4	16.43	8.34	3.70	47.25
B5	0.34	0.37	0.00	1.73

TABLE 4

Means, Standard Deviations and Ranges for Structural Variables

Variable	Mean	Standard Deviation	Low	High
S1	19.99	20.28	4.00	138.00
S2	11.64	6.19	1.00	31.00
S3	1.07	1.81	0.00	9.00
S4	1.03	0.91	0.00	5.00
S5	0.46	0.82	0.00	3.00
S6	0.22	0.42	0.00	1.00
S7	0.19	0.39	0.00	1.00
S8	0.24	0.59	0.00	3.00
S9	0.45	0.68	0.00	3.00
S10	0.07	0.25	0.00	3.00
S11	0.25	0.52	0.00	2.00
S12	0.09	0.37	0.00	2.00
S13	3.84	2.85	1.00	15.00
S14	15.13	3.90	5.00	19.00
S15	0.37	1.14	0.00	6.00
S16	1.28	1.91	0.00	5.00
S17	5.44	4.88	0.00	23.00

In comparing B1, Table 3, with S13, Table 4, we see that the students have been very successful in producing minimal proofs. However, this must be considered in light of the fact that there were 51 problems which could be solved by a one-line proof, a fact which was reflected in the behavioral data where there were a total of 47 problems for which the B1 value was less than 2.00. Further, the variance of B1 for some of the longer problems is quite large indicating that fewer students produced a minimal proof for these problems.

Tables 5 and 6 contain the problem statements and the standard proofs for the seven problems having extreme values on the behavioral measures. For the low values on variables B1 and B5, the problems were chosen arbitrarily from those with the appropriate magnitude. Tables 5 and 6 provide insight into the features of the problems and curriculum which give rise to extreme values on the behavioral measures. A familiarity with these logic problems will add meaning to the discussion of the analysis presented below. A brief explanation of each of these problems is given followed by a discussion of the relationships among the variables for these problems. Readers unfamiliar with the rules of LIS may refer to Appendix B.

Problem 415032 received a value of 15.80 for B1. It begins with a hint telling the student that there is a certain redundancy in the rules which he has available. At this point in the curriculum he has been given all five axioms for addition plus the first three theorems. He is asked to derive 6 = 3 + 3. It is possible to produce a derivation using the axioms and theorems, but, this will not result in the minimal proof. To obtain the minimal proof the student must use the rules learned earlier in the curriculum.

This problem requires 14 lines for its standard solution. In addition, it would be considered a difficult problem on all of the measures considered. It is ranked (15) on measure B2 with a latency of 252.91 sec., (20) on measure B5 with .9 error messages and (85) on measure B4 with 16.24 sec. per line. The problem involves five applications of the rule ND and appropriate algebraic manipulations and algebraic substitutions which are accomplished in this case by the rules AR, CE, CA, and RE.

Insert Table 5 about here

The problem ranked highest on measure B2 is 413010. Again this problem would be considered very difficult on all of the measures. It is ranked (3) on measure B5 with 1.57 error messages, (8) on measure B4 with 37.86 sec/line, and (12) on measure B1 with 11.09 lines. In this problem, the student is asked to derive the conditional: if A is less than (5+4)+1 then A is less than 5+((1+3)+1). The student can easily verify that it is true since obviously (5+4)+1 equals 5+((1+3)+1). One approach could be to show that $A < 10 \rightarrow A < 10$ and then show $10=(5+4)+1=5+((1+3)+1)$ and substitute. However, this would require more than seven lines. There are, of course, several other approaches.

The problem ranking highest on measure B5 is 414030. In this problem the student must derive the statement that A+(3+(-A)) equals 1+(1+1), a statement which is obviously true. This problem is similar to problem 415032 except that conditional proof is not required and the student, at this point in the curriculum, has no theorems available to him. As in the two previous cases, it would be considered a difficult problem on the other measures also. It was (2) on measure B1 with 15.68

TABLE 5

Problems Receiving Highest Value on Behavioral Variables

415.32:

THERE ARE SOME SUPERFLUITIES AMONG OUR RULES, BUT SINCE WE ARE NOT AFTER MATHEMATICAL ELEGANCE WE TOLERATE THEM. A SIMILAR SITUATION EXISTS IN THE RULES OF SENTENTIAL LOGIC.

DERIVE: $6 = 3 + 3$

ND6 (1) $6 = 5 + 1$
ND5 (2) $5 = 4 + 1$
1.2RE1 (3) $6 = (4 + 1) + 1$
3AR2 (4) $6 = 4 + (1 + 1)$
ND2 (5) $2 = 1 + 1$
5CE1 (6) $1 + 1 = 2$
4.6RE1 (7) $6 = 4 + 2$
ND4 (8) $4 = 3 + 1$
7.8RE1 (9) $6 = (3 + 1) + 2$
9AR2 (10) $6 = 3 + (1 + 2)$
ND3 (11) $3 = 2 + 1$
11CE1 (12) $2 + 1 = 3$
12CA1 (13) $1 + 2 = 3$
10.13RE1
(14) $6 = 3 + 3$

CORRECT

413.10:

DERIVE: $A < (5 + 4) + 1 \rightarrow A < 5 + ((1 + 3) + 1)$

WP (1) $A < (5 + 4) + 1$
AS $(A + B) + C = A + (B + C)$
A:5
B:4
C:1 (2) $(5 + 4) + 1 = 5 + (4 + 1)$

1.2RE1 (3) $A < 5 + (4 + 1)$

ND4 (4) $4 = 3 + 1$

3.4RE1 (5) $A < 5 + ((3 + 1) + 1)$

5CA2 (6) $A < 5 + ((1 + 3) + 1)$

1.6CP (7) $A < (5 + 4) + 1 \rightarrow A < 5 + ((1 + 3) + 1)$

CORRECT

405.23:

DERIVE: $A = C$

P (1) $A + B = A + C \rightarrow B = C$

P (2) $B = C \rightarrow A = C$

P (3) $A + B = A + C$

1.3AA (4) $B = C$

3.4AA (5) $A = C$

CORRECT

415.30:

DERIVE: $A + (3 + (-A)) = 1 + (1 + 1)$

AI $A + (-A) = 0$

A:A (1) $A + (-A) = 0$

1AE

:3 (2) $(A + (-A)) + 3 = 0 + 3$

2AR2 (3) $A + ((-A) + 3) = 0 + 3$

3CA2 (4) $A + (3 + (-A)) = 0 + 3$

Z $A + 0 = A$

A:3 (5) $3 + 0 = 3$

5CA1 (6) $0 + 3 = 3$

4.6RE1 (7) $A + (3 + (-A)) = 3$

ND3 (8) $3 = 2 + 1$

7.8RE2 (9) $A + (3 + (-A)) = 2 + 1$

ND2 (10) $2 = 1 + 1$

9.10RE1

(11) $A + (3 + (-A)) = (1 + 1) + 1$

11AR4 (12) $A + (3 + (-A)) = 1 + (1 + 1)$

CORRECT

lines, (5) on measure B2 with 303.17 sec. and (65) on measure B4 with 18.20 sec. per line. This problem also requires the derivation of a complex formula and involves the use of two axioms, AI and Z, as well as some complicated algebraic manipulations and substitutions.

Problem 405023, which ranked highest on measures B4 and B3, differs from the previous problems in two interesting ways. First, it does not rank very high on the other measures. For B2, it is (65) with a latency of 94.50. For B5, it is (117) with .07 error messages. For B1, it is (151) with 2.00 lines. This problem requires only two applications of rule AA for its solution and, thus, does not seem intuitively difficult. However, one might explain its observed difficulty by the fact that it was preceded by 19 multiple-choice problems. This problem offers a dramatic illustration of the effects of surrounding context on student performance on a particular problem.

Insert Table 6 about here

Table 6 contains those problems which received the lowest values on the behavioral measures. These problems have several features in common and thus they are discussed as a group. First, they are all problems which require only one line for their solution. Second, each problem would be rated as "easy" on all of the behavioral measures. Third, each problem has a value of zero on measure B5. For problems 412023 and 414004 the student is told exactly what he must type in order to obtain the solution. The slightly higher than minimal latencies for these problems are probably due to the time required for the student to read the accompanying text. Problem 414005 is of precisely the same

TABLE 6

Problems Receiving Lowest Values

on Behavioral Variables

412.23:

```
THERE IS A SHORT FORM OF CA SIMILAR IN SOME RESPECTS TO USES OF RE.
IN ORDER TO DERIVE A + B = 3 + 6 FROM THE PREMISE A + B = 6 + 3
SIMPLY TYPE '1CA2'.

DERIVE:  A + B = 3 + 6

P          (1)      A + B = 6 + 3

1CA2       (2)      A + B = 3 + 6

CORRECT
```

414.4:

```
TO USE THE Z AXIOM
          1)  TYPE 'Z' AND SPACE
          2)  AFTER THE COMPUTER TYPES IN THE AXIOM AND 'A:' TYPE
              THE TERM YOU WANT TO REPLACE 'A'.

DERIVE:  5 + 0 = 5

Z A + 0 = A
A:5       (1)      5 + 0 = 5

CORRECT
```

414.5:

```
DERIVE:  17 + 0 = 17

Z A + 0 = A

A:17      (1)     17 + 0 = 17
CORRECT
```

type as the preceding problem 414004, which introduces the Z axiom. The only difference is that the student must type 17 instead of 5 for the substitution into the axiom.

Tables 7, 8 and 9 contain the correlations of the behavioral variables with one another, the structural variables with one another and the behavioral variables with the structural variables, respectively. These correlations were obtained as part of the standard output of the BMD06M program. The results have been separated into three tables for ease of examination and discussion.

In Table 7, we find several interesting correlations which give some insight into the nature of the relationships among the various measures of difficulty. First, observe that B1 is highly correlated with B2 and B5 but not with B3 and B4. It is not surprising that latency and error rate increase with the length of proof. However, it is reassuring to see the correlation of B1 with B3 and B4 is not high, indicating that our measure of difficulty is not a simple function of problem length. The correlations between B2 and, B3 and B4, are seen to be somewhat higher. The almost perfect correlation of B3 and B4 provides further evidence that they are measuring the same underlying behavior and further justification for the decision to choose only one of them as the difficulty measure (B4). One final observation is that all of the correlations among our behavioral measures are positive and equal to or greater than .37.

Insert Table 7 about here

Table 8 contains the correlations of the structural variables with one another. Since these variables are defined solely in terms of

TABLE 7

Correlations of Behavioral Variable

	B1	B2	B3	B4	B5
B1	1.00	0.90	0.37	0.40	0.77
B2		1.00	0.64	0.68	0.88
B3			1.00	0.99	0.57
B4				1.00	0.61
B5					1.00

curriculum structure, an examination of their correlations will provide

---

Insert Table 8 about here

---

some insight into certain features of the curriculum. The majority of the correlations are low; only 22 of the 185 correlations have an absolute value greater than .40. The large number of low correlations is desirable because an attempt was made to define the variables so that they reflect nonredundant features of the curriculum. Since it is impractical to discuss each of the 185 correlations, only those variables which appear to be of most interest are discussed.

In examining the correlations, we are able to distinguish two patterns. First, a number of correlations are indicative of the 500 lessons. It should be recalled that these lessons deal with sentential logic. This is reflected in the high correlations between S6 and S3, S4, S9, S10, S14. For example, the high positive correlation between S3 and S6 indicates that a greater number of logical connectives are found in problems on logic than in problems on algebra. The correlation between S6 and S4 indicates more nesting of parentheses in the first part of the curriculum and the correlation between S6 and S9 and S6 and S10 suggest more frequent use of conditional proof (CP) and proof by contradiction (IP). The high correlation between S6 and S14 reflects the fact that most of the rules become available in the first part of the curriculum. This is further supported by the high positive correlations between S3 and S9, and S3 and S10. There is also evidence that the proofs are longer in the first part of the curriculum than in the second part because of the correlation between S3 and S13 and a correlation of 0.30 between S6 and S13.

TABLE 8

Correlations of Structural Variables

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17
S1	1.00	0.18	0.05	0.13	-0.12	0.12	0.63	-0.17	-0.01	-0.07	0.03	-0.10	-0.04	-0.07	-0.06	-0.03	-0.23
S2		1.00	0.12	0.62	-0.32	-0.06	-0.20	-0.14	0.14	-0.04	-0.06	-0.04	0.17	0.10	-0.15	-0.04	0.23
S3			1.00	0.52	-0.24	0.87	-0.10	0.27	0.80	0.52	-0.27	-0.15	0.43	-0.72	-0.19	0.37	0.00
S4				1.00	-0.43	0.49	-0.20	0.08	0.39	0.37	-0.04	0.04	0.26	0.30	-0.15	-0.09	0.00
S5					1.00	-0.26	-0.04	0.48	-0.23	-0.15	0.24	-0.14	-0.10	0.04	-0.15	0.26	0.10
S6						1.00	-0.01	0.12	0.74	0.51	0.26	-0.14	0.30	0.83	-0.17	0.36	-0.09
S7							1.00	-0.20	-0.17	-0.13	-0.13	-0.02	-0.17	0.02	0.20	0.19	0.34
S8								1.00	0.22	0.09	-0.20	-0.10	0.25	-0.32	-0.13	0.27	0.10
S9									1.00	0.37	0.28	-0.17	0.45	-0.65	-0.20	0.31	0.10
S10										1.00	-0.13	-0.07	0.25	0.27	-0.09	-0.18	0.03
S11											1.00	0.01	0.05	0.44	0.19	0.67	-0.17
S12												1.00	-0.02	0.25	0.73	0.50	-0.15
S13													1.00	-0.13	0.05	0.03	0.23
S14														1.00	0.32	0.65	0.08
S15															1.00	0.63	0.14
S16																1.00	0.23
S17																	1.00

The second pattern consists of those high correlations which arise as a result of the manner in which the variables are defined. For example, there is a positive correlation between S11 and S16 because an axiom cannot occur in a proof before the number of axioms available become greater than zero. Similar explanations, based on the definition of the variables, can be given for the correlations between S12 and S15, S12 and S16, S14 and S16, S15 and S16, S1 and S7, S2 and S4 and S3 and S4.

Finally, two other correlations which appear in the analysis are worth mentioning. First, there is a correlation of .48 between S5 and S8. It appears that problems which use several occurrences of AA have the greatest number of premises. An example of such a problem is problem 405023 in Table 5. Second, the high correlation between S9 and S13 indicates that problems requiring conditional proof tend to be longer than those not requiring the use of this rule.

The discussion now turns to an examination of the relationship between the two sets of variables. The correlation between the behavioral and structural variables can be found in Table 9. Next, the relationship is described more formally by means of the canonical correlation analysis. Finally, the predictive models obtained from the regression analyses are presented, first the models which have variables B1 and B2 as the dependent variable and then in more detail, the model in which difficulty (variable B4) is the dependent variable.

Insert Table 9 about here

The correlations found in Table 9 between the two sets of variables are rather low and in the majority of cases almost zero. The largest

TABLE 9

Correlations Between Behavioral and Structural Variables

	B2	B3	B4	B5	B6
S1	-0.06	-0.03	0.05	0.05	0.01
S2	0.21	0.16	-0.03	-0.03	0.10
S3	0.44	0.37	0.17	0.18	0.25
S4	0.34	0.37	0.11	0.13	0.25
S5	-0.16	-0.08	0.24	0.27	-0.09
S6	0.36	0.34	0.19	0.23	0.27
S7	-0.19	-0.13	0.07	0.07	-0.11
S8	0.17	0.11	0.06	0.05	0.05
S9	0.44	0.34	0.09	0.11	0.25
S10	0.33	0.31	0.23	0.27	0.30
S11	0.06	0.08	-0.03	-0.02	0.15
S12	0.05	0.05	0.05	0.04	0.04
S13	0.93	0.74	0.27	0.27	0.60
S14	-0.13	-0.10	-0.07	0.08	-0.06
S15	0.09	0.08	0.09	0.07	0.07
S16	0.08	0.08	0.03	0.01	0.09
S17	0.19	0.10	-0.03	-0.03	0.05

correlations are those between S13 and the behavioral variables B1, B2 and B5. Also there are high correlations between the minimal number of lines in a proof and the actual length, latency and number of error messages for the proof.

Variables S3, S6 and S9 are also highly correlated with B1, B2 and B5. However, as is evident from Table 4, these structural variables are also very highly correlated with each other and it is not easy to interpret their effect on the behavioral variables from Table 9 alone. Variable S10 also appears to be important. This variable is discussed in more detail later.

The structural variables most highly correlated with the difficulty variable B4 are S5, S10 and S13, all .27. From Table 4, it can be seen that these structural variables are not highly correlated with each other. They play an important role in the regression model discussed below. Note that most of the remaining structural variables have almost zero correlations with B4. Thus, we are led to consider models which involve linear combinations of the variables.

Table 10 contains the results of the canonical analysis.

Insert Table 10 about here

Behavioral variable B3 is omitted from the analysis for the reasons discussed in Chapter II and above. Thus, there were four canonical correlations and four sets of coefficients for the canonical variates. Since I am interested only in describing the dependencies among the variables and do not intend to use the derived variates for later analyses, I have not explicitly computed the canonical variates from

TABLE 10

Canonical Correlations and Coefficients

Canonical Correlation = 0.94682261

Coefficients for the first set of variables:

-1.326259(B1)	0.284225(B2)	0.021786(B4)	0.107016(B5)

Coefficients for the second set of variables:

0.073856(S1)	-0.146555(S2)	0.123607(S3)	0.057187(S4)
0.047831(S5)	-0.269110(S6)	0.009731(S7)	-0.018433(S8)
0.005036(S9)	-0.036156(S10)	0.050815(S11)	-0.035614(S12)
-0.933880(S13)	-0.048610(S14)	0.007074(S15)	-0.084335(S16)
0.006210(S17)			

Canonical Correlation = 0.52323435

Coefficients for the first set of variables:

-0.089763(B1)	0.224815(B2)	-1.261107(B4)	0.335463(B5)

Coefficients for the second set of variables:

0.101760(B1)	-0.290956(S2)	0.164078(S3)	-0.096213(S4)
-0.157033(S5)	-0.951645(S6)	-0.321546(S7)	0.324336(S8)
-0.0[illegible]440(S9)	-0.313280(S10)	0.029467(S11)	-0.020757(S12)
0.093578(S13)	-0.322343(S14)	-0.196313(S15)	-0.139318(S16)
0.048549(S17)			

Canonical Correlation = 0.37973930

Coefficients for the first set of variables:

1.844434(B1)	-1.473808(B2)	0.818170(B4)	-1.291125(B5)

Coefficients for the second set of variables:

-0.193068(S1)	0.060709(S2)	1.114298(S3)	-0.114576(B4)
0.051367(S5)	-0.928428(S6)	0.366820(S7)	0.176483(S8)
-0.011327(S9)	-0.387504(S10)	-0.688379(S11)	-0.125276(S12)
-0.177405(S13)	-0.464047(S14)	-0.006174(S15)	0.685214(S16)
0.401779(S17)			

Canonical Correlation = 0.22914162

Coefficients for the first set of variables:

1.913588(B1)   -4.001316(B2)   0.961281(B4)   1.651518(B5)

Coefficients for the second set of variables:

0.518413(S1)   -0.550062(S2)   -1.176869(S3)   0.189711(S4)
-0.054340(S5)   1.058815(S6)   -0.120565(S7)   0.280015(S8)
0.500092(S9)   -0.461083(S10)   0.125329(S11)   -0.316266(S12)
0.039960(S13)   0.477343(S14)   0.603351(S15)   -0.321938(S16)
0.356688(S17)

the coefficients. In the table, the canonical correlation is followed first by the set of coefficients for the behavioral variables, namely, B1, B2, B4 and B5, and then the coefficients for the structural variables, S1 through S17. In interpreting the coefficients in Table 6, one must remember that the canonical correlations were obtained from the covariance matrix. Thus, the magnitude of the coefficients depends on the magnitude of the variables considered. To illustrate what this means, consider variables B2 and B5 and their respective coefficients for the canonical correlation 0.52. From Table 3, we see that the mean for B2 is 84.18 and the mean for B5 is .34; the coefficients are .22 and .34 for B2 and B5, respectively. Thus, on the average, B2 contributes 18.52 units to the canonical variate whereas B5 contributes only 0.77. Ignoring the magnitudes of the variables, one would say that variable B5 plays the more important role due to the larger magnitude of its coefficient but when the magnitudes of the contribution are considered, it is B2 which makes, by far, the larger contribution to the canonical variate.

For the maximum canonical correlation .95, the canonical variate for the behavioral variables places the most weight on B1 and B2. The canonical variate for the structural variables places the most weight on S1, S2 and S13. Essentially, the first variate is some measure of the length of a problem, that is, a linear combination of number of lines and latency. Similarly, its correlative in the concomitant variables is a structural measure of length, where S1 and S2 are measures of the amount of information to be processed and S13 is the minimal length of a proof. Thus, the first correlation establishes a link between the behavioral measures of length of a problem and their structural counterparts.

The magnitude of the correlation indicates that the relationship between these variables is a very strong one.

The second canonical correlation 0.52 appears to place the greatest weight on variables B2 and B4 for the behavioral variate and on variables S1, S2, and S14 for the structural variate. This case yields, primarily, a comparison of "difficulty" expressed as a weighted sum of B2 and B4 with "structural complexity" expressed as a weighted sum of S1 and S2, information to be processed, and S14, availability of rules. The variable S14 appears to make the greatest contribution to the structural canonical variate.

The final two canonical correlations are rather low and, thus, their corresponding derived variates are not of as much interest as those described above. For both of these correlations, the most important structural variables are S1 and S14. In addition, for the 0.38 correlation, variable S17 contributes heavily to the structural variate and for the 0.23 correlation, variable S2 is the other heavily weighted variable.

The procedure used for the regression analyses is considered next. Using the results of the canonical correlation analysis as a guide, I ran three separate regression analyses in which B1, B2 and B4 were the dependent variables. The plots described in Chapter III, p. 28 were obtained as part of the output for these regressions. An examination of these plots reveals that variables B2 and B4 appear to violate the homoscedasticity assumption. After applying a square-root transformation to variables B2 and B4, we find that this assumption appears to be satisfied.

For example, Figure 1 shows the plot of the residuals versus variable B4. One can observe a rather obvious dependence of magnitude of residuals on magnitude of B4 (see dotted lines). In Figure 2, the same plot is shown after applying the square-root transformation to B4. Notice that the pattern, which was observed in Figure 1, no longer appears.

Insert Figures 1 and 2 about here

Several transformations were applied to some of the independent variables also. However, none of the transformed variables, except for the cube of S5, entered into the regression equations.

The regressions were redone, this time using variables B1, $\sqrt{B2}$ and $\sqrt{B4}$ as the dependent variables. The results for these regressions may be found in Tables 11, 12, and 13. These tables give the step at which each variable entered the regression, the value of R and R^2 at that step, the increase in R^2 due to the addition of that variable, the F-value required for deletion and the final regression coefficient for the variable. It would be pointless to discuss any variable which did not contribute at least 1 percent to R^2 and such variables have been eliminated from the models. The Anova tables are given only for the actual models used. They contain the variables in the equation with the step that the variable entered, the coefficient, the standard error of the coefficient and its computed t-value, the multiple correlation coefficient, and the standard error of estimate of Y.

Insert Tables 11, 12, 13 about here

Table 14 contains the results for variable B1. Variable S13 accounts for 86 percent of the variation in this case. Since S13 is the

Insert Table 14, about here

PLOT: RESIDUALS(Y-AXIS) VS COMPUTED Y (X-AXIS)

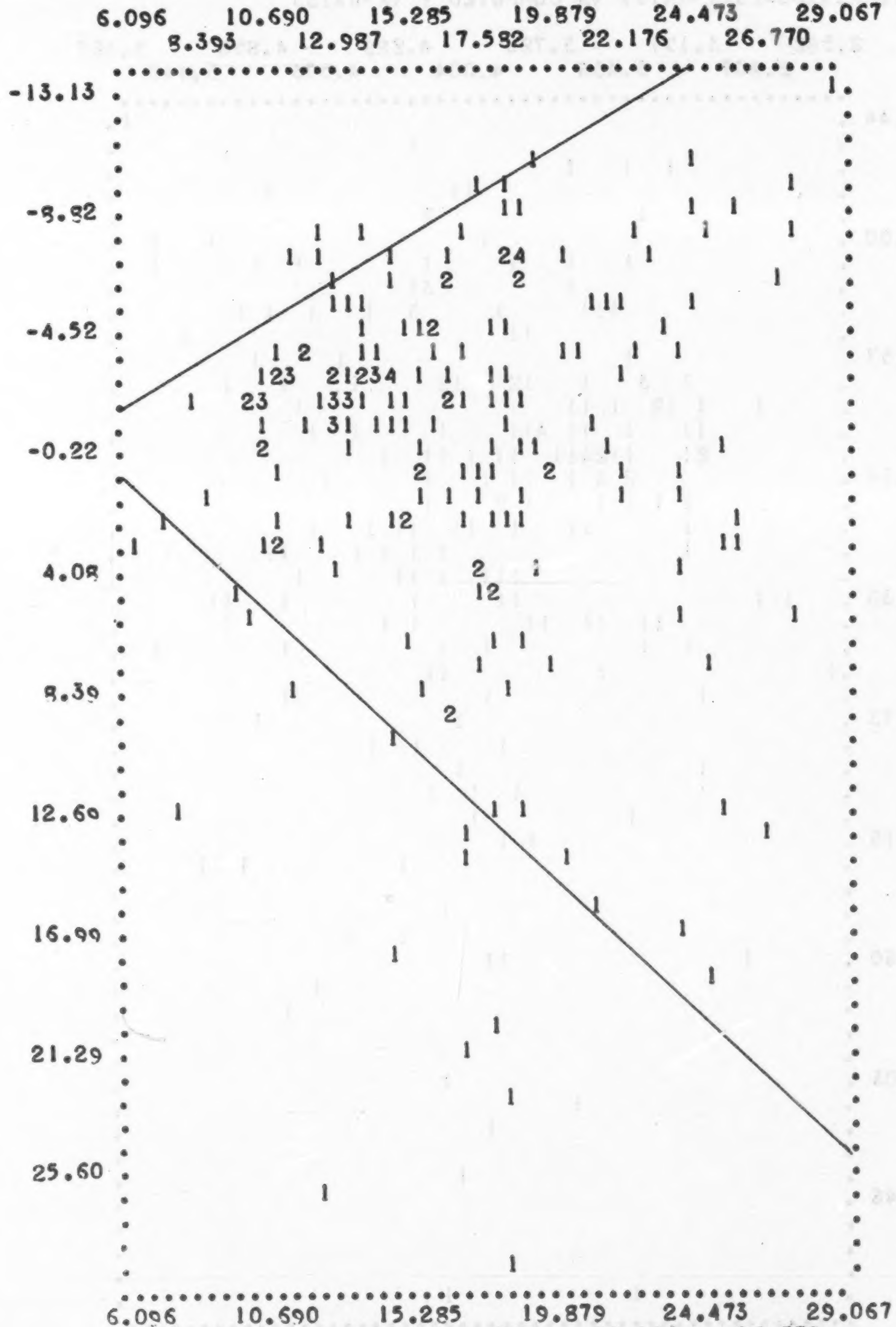

Figure 1

PLOT: RESIDUALS(Y-AXIS) VS COMPUTED Y (X-AXIS)

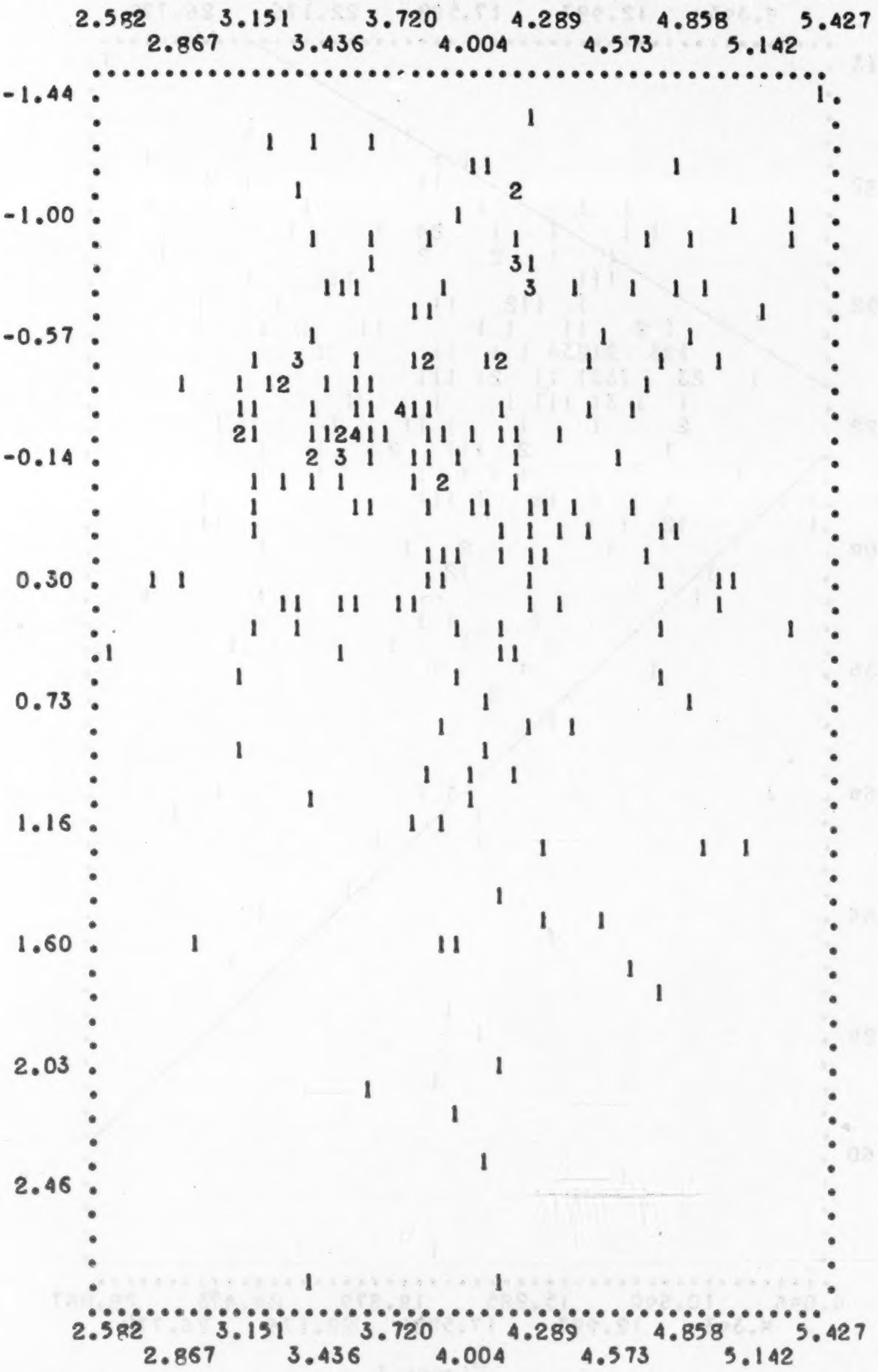

Figure 2

TABLE 11

Summary Table for Variable B1

Step Num.	Variable Ent. Rem.	Multiple R	R^2	Increase in R^2	F Value For Del.	Last Reg. Coefficients
1	S13	0.93150	0.86769	0.86769	1318.6469	1.16334
2	S12	0.93590	0.87591	0.00822	13.0515	1.16625
3	S10	0.93900	0.88172	0.00581	9.7616	0.90045
4	S8	0.94080	0.88510	0.00338	5.9794	-0.28200
5	S15	0.94170	0.88680	0.00169	3.0351	-0.23345
6	S16	0.94280	0.88887	0.00207	3.6043	0.06910
7	S6	0.94330	0.88981	0.00094	1.7474	2.69823
8	S3	0.94520	0.89340	0.00359	6.4720	-0.36914
9	S2	0.94700	0.89681	0.00341	6.1304	0.07617
10	S5	0.94850	0.89965	0.00284	5.4828	0.28674
11	S1	0.95000	0.90250	0.00285	5.7656	-0.01312
12	S4	0.95030	0.90307	0.00057	1.0777	-0.18570
13	S14	0.95060	0.90364	0.00057	1.0770	0.07470
14	S17	0.95090	0.90421	0.00057	1.0968	-0.01728
15	S7	0.95100	0.90440	0.00019	0.3109	0.17845
16	S11	0.95100	0.90440	0.00000	0.1903	0.09323

TABLE 12

Summary Table for Variable $\sqrt{B2}$

Step Num.	Variable Ent. Rem.	Multiple R	Multiple R^2	Increase In R^2	F Value For Del.	Last Reg. Coefficients
1	S13	0.78690	0.61921	0.61921	326.9081	1.17775
2	S10	0.80120	0.64192	0.02261	12.6375	2.30444
3	S12	0.81050	0.65691	0.01499	8.7281	1.70137
4	S18	0.81860	0.67011	0.01320	7.9325	0.09758
5	S8	0.82650	0.68310	0.01300	8.1053	-0.80486
6	S6	0.83220	0.69256	0.00945	5.9470	4.87922
7	S11	0.83750	0.70141	0.00885	5.8126	0.62469
8	S2	0.84120	0.70762	0.00621	4.1587	0.12205
9	S5	0.84510	0.71419	0.00658	4.4065	0.88308
10	S3	0.84770	0.71860	0.00440	2.9721	-0.47333
11	S14	0.84970	0.72199	0.00339	2.3257	0.15648
12	S1	0.85030	0.72301	0.00102	0.7322	-0.01469
13	S17	0.85110	0.72437	0.00136	0.8755	-0.03637
14	S7	0.85150	0.72505	0.00068	0.5108	0.43606
15	S9	0.85170	0.72539	0.00034	0.2904	0.22169
16	S4	0.85180	0.72556	0.00017	0.0933	-0.12122
17	S16	0.85190	0.72573	0.00017	0.0432	0.05144
18	S15	0.85190	0.72573	0.00000	0.0330	-0.04807

TABLE 13

Summary Table for Variable $\sqrt{B4}$

Step Num.	Variable Ent. Rem.	Multiple		Increase In R^2	F Value For Del.	Last Reg. Coefficients
		R	R^2			
1	S13	0.27500	0.07563	0.07563	16.4435	0.09061
2	S5	0.36900	0.13616	0.06054	14.0246	0.68067
3	S10	0.43250	0.18706	0.05090	12.4530	0.66759
4	S8	0.47630	0.22686	0.03981	10.1854	-0.30851
5	S6	0.50680	0.25685	0.02998	7.9536	1.35355
6	S16	0.54620	0.29833	0.04149	11.5861	0.02457
7	S2	0.56010	0.31371	0.01537	4.3769	0.02947
8	S7	0.57170	0.32684	0.01313	3.7903	0.37023
9	S12	0.57690	0.33281	0.00597	1.7321	0.18215
10	S3	0.58000	0.33640	0.00359	1.0083	-0.09061
11	S17	0.58230	0.33907	0.00267	0.7875	-0.01447
12	S14	0.58480	0.34199	0.00292	0.8292	0.04600
13	S15	0.58620	0.34363	0.00162	0.5003	0.06032
14	S1	0.58700	0.34457	0.00094	0.2493	-0.00204
15	S11	0.58740	0.34504	0.00047	0.1462	0.06531
16	S4	0.58760	0.34527	0.00023	0.0553	0.03150
17	S9	0.58760	0.34527	0.00000	0.0193	0.02255

TABLE 14

ANOVA Table and Significant Variables for B1

Analysis of Variance:

	DF	Sum of Squares	Mean Square	F-Ratio
Regression	1	1935.42	1935.42	1392.39
Residual	201	279.20	1.39	

Variables in Equation: (Constant = .295)

Variable	Step Entered	Coefficient	Std. Error	Computed T-Value
S13	1	1.11	.03	37.00*

Number of steps	1
Multiple R	0.93
Multiple R^2	0.87
Std. Error of Est.	0.10

*p < .001

number of lines in the minimal proof, one can say, with the qualifications mentioned on p. 33, that the students were quite successful in finding the minimal proofs. The remaining variables account for only an additional 4 percent increase in R^2. Thus, it appears that the more interesting aspects of performance on the logic problems are not reflected in the problem length.

Table 15 contains the results for the regression using the square root of total latency, $\sqrt{B2}$, as the dependent variable. In this case,

Insert Table 15 about here

the model was able to account for 68 percent of the variation in total latency with six variables. The value for R^2 is significantly nonzero at $p < .01$.

The most important variable and the first to enter the equation is variable S13, the number of lines in the minimal proof. It is not surprising that the amount of time spent on a problem is very strongly dependent on its length. However, the other variables included in this model begin to give insight into some of the other factors affecting the time a student spends on a problem.

The second variable to enter the equation is variable S10, the number of occurrences of IP in the standard proof. The increase in latency may be attributed to two factors. First, the rule requiring three arguments, is complicated to use; the error rate for problems requiring the use of the rule IP was, in general, higher than for other problems. Second, a student must spend time to discover the contradiction needed for the indirect proof.

TABLE 15

ANOVA Table and Significant Variables for the Square-Root of B2

Analysis of Variance:

	DF	Sum of Squares	Mean Square	F-Ratio
Regression	6	2556.52	426.09	73.57
Residual	196	1135.17	5.97	

Variables in Equation: (Constant = 2.99)

Variable	Step Entered	Coefficient	Std. Error	Computed T-Value
S6	6	1.22	0.50	4.18**
S8	5	-1.06	0.34	2.82*
S10	2	2.27	0.78	2.91*
S12	3	1.63	0.47	3.47**
S13	1	1.21	0.07	17.28**
S18	4	0.15	0.04	3.75**

Number of Steps	6
Multiple R	0.83
Multiple R^2	0.68
Std. Error of Est.	2.41

*$p < .01$
**$p < .001$

The third significant variable to enter the equation is S12, the number of occurrences of a theorem in the minimal proof. The increase in latency due to the presence of theorems in a proof may be explained as follows. Unlike rules and axioms, there are no mnemonics for the theorems. If a student feels that a theorem is appropriate, he must first consult his theorem sheet to see if there is such a theorem and to find its number (e.g., TH3). Thus, except in the improbable event that a student has memorized the theorem numbers, these problems require more time, even though they are not necessarily more difficult.

The transformed variable S18, the cube of the number of premises, enters the equation next. This variable represents, in part, the information to be processed by the student before he solves the problem. Each additional premise greatly increases the amount of time spent on the problem.

The fifth significant variable to enter the equation is S8, the number of occurrences of AA in the minimal proof. Note that this variable has a negative coefficient. This variable was also significant in the regression equation obtained for $\sqrt{B4}$, where it also received a negative coefficient. An interpretation for it is given in the discussion below.

The final variable in the model for latency is S6, the problem context. This variable indicates that, on the average, the problems in the CEX portion of the curriculum require more time.

None of the remaining variables contribute as much as 1 percent to R^2, as can be seen from Table 12. Thus, they are not included in the model for latency.

Table 16 contains the results of the regression which used $\sqrt{B4}$, square root of latency per line, as the dependent variable. Those

---

Insert Table 16 about here

---

variables which contribute over 1 percent to R^2 and are significantly nonzero were chosen for the model. With the seven variables meeting this criterion, the model was able to account for 33 percent of the variation. Although this value for R^2 is not as impressive as the values in the previous two cases, the F-ratio of 12.735 is significant for $p < .01$. Further, an examination of the important variables in this first attempt to predict problem difficulty has revealed some of the important structural features which may be further broken down and explored in future studies of this nature. Some possibilities are considered in Chapter V. But first, the results of the present analysis are presented.

Variable S13, the number of lines in the standard proof, is the first variable to enter the equation. It accounts for 8 percent (see Table 13) of the total variation. Thus, the length of a proof is an indicator of difficulty, but it does not assume the overwhelming importance which it had in the two previously discussed models.

The second variable to enter is S5, the number of premises, and it accounts for an additional 6 percent of the variation. The great majority of problems in which premises are given are to be found in the CEX portion of the curriculum. Hence, this variable may also be accounting for part of the effect due to problem context along with the information to be processed.

Variable S10, the number of occurrences of IP in the standard proof, which accounted for an additional 5 percent of the variation, enters the

TABLE 16

ANOVA Table and Significant Variables for the Square-root of B4

Analysis of Variance:

	DF	Sum of Squares	Mean Square	F-Ratio
Regression	7	58.62	8.37	12.74
Residual	195	128.23	0.66	

Variables in Equation: (Constant = 2.66)

Variable	Step Entered	Coefficient	Std. Error	Computed T-Value
S2	7	0.02	0.01	2.20*
S5	2	0.68	0.09	7.56**
S6	5	0.87	0.19	4.69**
S8	4	-0.40	0.12	3.33*
S10	3	0.70	0.26	2.69*
S13	1	0.07	0.03	2.33*
S16	6	0.14	0.04	3.50*

Number of Steps	7
Multiple R	0.56
Multiple R^2	0.33
Std. Error of Est.	0.81

* p .01

** p .001

equation next. In addition to the extra time required to use this rule (see p. 57, a problem involving the use of IP requires a different kind of behavior on the part of the student than that required in a straight derivation problem. The results imply that this difference is significant and results in increased difficulty.

The only variable to have a negative coefficient is variable S8, the number of occurrences of AA in the standard proof. This variable accounts for 4 percent of the variation. Table 9 shows that this variable is highly correlated with S5, thus making it somewhat difficult to interpret. Note further that the AA rule was used predominantly in the CEX portion of the curriculum and only in those problems which could not be solved by means of a counterexample. That is, AA appeared only in DERIVE-type problems. Thus, this variable might be interpreted as accounting for the fact that in context of the CEX portion of the curriculum, derive problems are easier than CEX problems.

Variable S6, the fifth variable to enter the regression equation, receives the largest coefficient. This is further evidence that problems in the CEX portion of the curriculum were more difficult than those in the remainder of the curriculum.

The sixth significant variable to enter is S16, the number of axioms available to the student. This variable gives a measure of the amount of information at the disposal of the student. This is the only case in which one of the "availability" variables (S14-S16) played a significant role.

Finally, the last significant variable to enter the regression equation is S2, the number of words in the sentence to be derived.

This variable is another measure of the information which must be processed by the student.

Seven significant variables which account for 33 percent of the variation in problem difficulty are identified. The first two, S2, the number of words in the sentence to be derived, and S5, the number of premises, are measures of the amount of information which must be processed by the student in order to solve the problem. S6 specifies whether a problem is included in the CEX part of the curriculum. The next three, S8, S10 and S13, are standard proof variables and reflect the nature of the required derivation. The final significant variable is S16, a measure of the amount of information available to the student, in this case the number of axioms.

In the next chapter, the results presented above are discussed. The discussion includes some of the implications and a possible extension of regression model.

CHAPTER V

DISCUSSION

The investigation described in the previous chapters was the first attempt to examine college student performance on LIS. In this chapter, we first comment upon the significant variables in the predictive difficulty model and define several new variables suggested by the results. Next we mention some of the other important results of our analysis and discuss the possibility of extending the regression model to a process or automaton model.

For purposes of the ensuing discussion, the seven significant variables are categorized under four major headings. The first category is problem context containing variable S6. The next category contains variables S2 and S5, which reflect the information which the student must process. The third category comprises three variables, namely, the standard proof variables S10, S8 and S13. The final category provides a measure of the available information with S16. One may write the predictive model as follows:

$$\sqrt{B4} = .87S6 + .02S2 + .68S5 - .40S8 + .70S10 + .07S13 + .14S16.$$

First consider problem context. The results show, without doubt, that the location of a problem in the curriculum is important. If a problem is in the CEX portion of the curriculum it is more difficult. In order to explore further the effect of a problem's position in the curriculum, I ran two additional regression analyses. In one analysis the dependent variable was $\sqrt{B4}$ for the 45 problems in the CEX portion of the curriculum, in the other analysis the dependent variable was $\sqrt{B4}$

for the remaining 158 problems. These analyses did not provide any additional information on the important features which predict problem difficulty. Thus, the procedure of grouping the two parts of the curriculum together did not adversely affect the results or mask the effect of any important variable.

It would also be of interest to determine if there is a sequential effect. If a sequential effect exists, the difficulty of a problem would be affected by the nature of the immediately preceding problem. In other words, if a DERIVE problem is more difficult when preceded by a CEX problem than when preceded by another DERIVE problem, we say there is a sequential effect. Define a (0,1) variable N1[*] which takes the value one if the preceding problem is of a different type and zero otherwise.

The second category deals with the information to be processed. Although five variables, S1-S5, have already been defined to provide a measure of this aspect of the problem, only two of them, S2 and S5, are significant in our model. Variable S2 is the number of symbols in the sentence to be derived. Although this is very crude measure, the variable is significant in predicting difficulty. A more refined measure of the information in the sentence to be derived would be of great value. However, the manner in which this information might be quantized is by no means obvious. As a step in the direction of capturing some of the information in the sentence to be derived, consider the following variable, N2, which retains the information provided by S2 while providing additional information about the sentence. Assign parentheses a base value of zero, all sentence letters, variables and constants a base value of one, unary operators a base value of two and binary operators a base value of three. Then the value of a symbol is its base value times the depth of nesting where we define the depth of nesting as S4 + 1. The value of N2 is the sum of the values of all of the symbols in the sentence to be derived. The following example is provided to illustrate N2. Suppose the problem is:

---

[*] Technically, N1 is a standard proof variable.

$$\begin{array}{rcccccccccccccccccccccccc}
 & 1 & 3 & 0 & 1 & 3 & 1 & 0 & 3 & 1 & 3 & 1 & 3 & 0 & 1 & 3 & 0 & 0 & 1 & 3 & 1 & 0 & 3 & 1 & 0 \\
\text{DERIVE:} & A & < & (& 5 & + & 4 &) & + & 1 & \rightarrow & A & < & (& 5 & + & (& (& 1 & + & 3 &) & + & 1 &) \\
 & 1 & 3 & 0 & 2 & 6 & 2 & 0 & 3 & 1 & 3 & 1 & 3 & 0 & 2 & 6 & 0 & 0 & 3 & 9 & 3 & 0 & 6 & 2 & 0
\end{array}$$

The number above the sentence are the base values of the symbols, the numbers below are the actual values. Their sum is 56, thus the value of N2 is 56. In future studies of this nature, more energy must be spent in trying to characterize the information in the sentence to be derived.

The second significant variable in this category is S5, the number of premises. As mentioned previously, since premises occur chiefly in the CEX portion of the curriculum, this variable may reflect, in part, the effect of problem context. In any case, the occurrence of premises in a problem does result in a considerable increase in difficulty and some of this increase is certainly due to the additional amount of information to be processed. Since the results indicate that premises are important, it would be of value to try to obtain a deeper understanding of the effect of premises. To do this we propose two new variables, N3 and N4. If there are no premises, N3 and N4 are zero. Before proceeding, one must distinguish between relevant and irrelevant premises. An irrelevant premise is one which is not used in the solution of the problem. With this distinction in mind, define N3 as the sum of the N2-values of each of the relevant premises and N4 as the sum of the N2-values of each of the irrelevant premises. These two new variables determine the effect of relevant and irrelevant premises on difficulty. They also provide a measure of the complexity of the premises.

The third category contains variables relfecting the nature of the required derivation, namely, the standard proof variables. In the model the three significant standard proof variables are S8, S10 and S13. The most important variable throughout the analysis has been S13, the number of lines in the standard proof. The other two significant standard proof variables involve the number of occurrences of specific rules in the standard proof, namely, AA and IP. Thus, one is led to consider trying other variables which reflect the nature of the required rules in a derivation, without going to the obviously impractical extreme of a separate variable for each rule. Define N5 as the number of different rules used in the standard derivation. Second, define N6n as the number of rules in the standard proof which require n arguments. This variable was suggested by the importance of variable S10.

The final category contains one significant variable, S16, the number of axioms available to the student. This variable provides some measure of the amount of information which the student has available to solve the problems. This variable brings to mind another issue, namely, the effect that "learning" a rule has on difficulty. For example, would variable S16 be significant if there had been data on a much more extensive portion of the curriculum, that is, if the study included all of the theorems on addition? By that time, presumably the axioms would have been well "learned" and perhaps variable S16 would no longer be of importance. At the present juncture in the research on student performance on logic problems, one may reasonably relegate such considerations to the status of "second-order" effects, but in the more refined stages of analysis they must be seriously considerd.

The following is a list of the suggested new structural variables:

N1	Sequential variable (0,1). Takes the value one if the preceding problem is of a different type, zero otherwise.
N2	Measure of complexity of sentence to be derived.
N3	Measure of the complexity of relevant premises.
N4	Measure of the complexity of irrelevant premises.
N5	Number of different rules used in derivation.
N6n	Number of rules in the standard proof requiring n arguments.

In addition to providing some first insights into the factors affecting problem difficulty, the present study yielded several other valuable results. First, the study resulted in a precise and intuitively satisfying definition of problem difficulty and provided a method of measuring it in terms of student protocols. Second, a large data base of student performance in elementary mathematical logic has been established from which it is possible to extract much more detailed information. It is hoped that other researchers and those interested in the teaching of logic will make use of this data base to further their understanding of student performance.

The effort to understand problem solving in mathematical logic should not stop with regression models. Suppes (1969) pointed out that "the main conceptual weakness of the regression models is that they do not provide an explicit temporal analysis of the steps being taken by a student in solving a problem." He then gave an example from research on arithmetic performance of elementary-school students which illustrates how an automaton model provides a natural tool for the analysis of data in arithmetic-problem solving.

Any mature theory of problem solving must account for the temporal sequence which a student goes through in solving a problem. That is, it must provide meaningful dynamic links of the variables which affect problem difficulty, variables such as those identified in this study. An automaton model would appear to be one of the more interesting possibilities for this purpose. Since all automata are, at least theoretically, programmable on a computer, the terms "automata" and "computer" will be used interchangably in the sequel.

The development of such models is possible, but the form that they should take is not yet clear. At present, there exist a number of computer programs which are able to prove theorems, i.e., solve problems such as those in the curriculum we have studied. However, the problems involved in developing the models are quite serious. First, we must find a theorem prover which "solves" problems in a manner analogous to the logic student. For example, a theorem prover based on the resolution principle (Robinson, 1965) is not appropriate. Then to analyze the student data, we must go from a deterministic model to a probabilistic one, that is, we must parameterize the model in such a way that it provides a good account of the performance data. In the case of arithmetic problems the structural variables identified in the regression models were of great value in parameterizing the atuomaton model.

I fully realize that the results presented here are still a long way from providing a fully adequate account of problem-solving in mathematical logic. I do, however, feel that I have taken an important initial step and gained some valuable insights into factors involved in performance on logic problems.

REFERENCES

Dixon, W. J. (Ed.) *Biomedical computer programs*. Berkeley: University of California Press, 1970.

Draper, N. R. and Smith, H. *Applied regression analysis*. New York: Wiley, 1966.

Hays, W. L. *Statistics for psychologists*. New York: Holt, Rinehart and Winston, 1963.

Kruskal, J. B. Statistical analysis: Transformations of data. International *Encyclopedia of the Social Sciences*. Vol. *15*. New York: McMillan and Free Press, 1968.

Morrison, D. F. *Multivariate statistical methods*. New York: McGraw-Hill 1967.

Robinson, J. A. A machine operated logic based on the resolution principle. *Journal of the Association for Computing Machinery*, 1965, *12*, 23-41.

Suppes, P. The uses of computers in education. *Scientific American*, 1966, *215*, 207-220.

Suppes, P. Computer-assisted instruction at Stanford. Technical Report 174, May 19, 1971. Stanford University, Institute for Mathematical Studies in the Social Sciences.

Suppes, P. and Binford, F. Experimental teaching of mathematical logic in the elementary school. *The Arithmetic Teacher*, 1965, *12*, 187-195.

Suppes, P. and Ihrke, C. Accelerated program in elementary school mathematics--The fourth year. *Psychology in the Schools*, 1970, *7* 111-126.

Suppes, P., Jerman, M. and Brian, D. *Computer-assisted instruction: Stanford's 1965-66 arithmetic program*. New York: Academic Press, 1968.

Suppes, P., Loftus, E. F. and Jerman, M. Problem-solving on a computer-based teletype. Educational Studies in Mathematics, 1969, 2, 1-15.

Wexler, J. D. A generative, remedial and query system for teaching by computer. Unpublished Ph.D. Dissertation, University of Wisconsin, 1970.

APPENDIX A

Examples of the different problem modes on the LIS

I. Multiple choice:

502.1:

IN THIS LESSON YOU WILL LEARN HOW TRUTH OR FALSITY OF A COMPLEX FORMULA IS RELATED TO THE TRUTH OR FALSITY OF ITS SUBFORMULAE. FOR INSTANCE, IF S IS FALSE, THEN ¬ S IS ...

A) TRUE
B) NOT TRUE

A

CORRECT

502.2:

IF S IS TRUE, THEN ¬ S IS ...

A) TRUE
B) NOT TRUE

B

CORRECT

II. Truth analysis mode:

503.23:

FIND THE TRUTH VALUES OF THE FOLLOWING:
S&(R&S)

T S
F R

LET'S COMPUTE THE TRUTH VALUES FOR ALL SUBFORMULAE
AND THEN FOR THE FORMULA ITSELF

F R&S
F S&(R&S)

CORRECT

502.36:

FIND THE TRUTH VALUE OF THE FOLLOWING:

```
¬ (¬ QV ¬ R)V ¬ (Q&R)

T    Q
T    R

     LET'S COMPUTE THE TRUTH VALUES FOR ALL SUBFORMULAE
     AND THEN FOR THE FORMULA ITSELF

F    ¬ Q
F    ¬ R
F    ¬ QV ¬ R
T    ¬(¬ QV ¬ R)
T    Q&R
F    ¬ (Q&R)
T    ¬ (¬ QV ¬ R)V ¬ (Q&R)

CORRECT
```

III. Counterexample mode:

505.17:

```
FIND A COUNTEREXAMPLE:       Q&S → R
P              (1)           QVR
P              (2)           R → (Q → S)
P              (3)           (S → R) → Q
CEX            YOU ARE LOOKING FOR A COUNTEREXAMPLE.  THUS, THE
               CONCLUSION MUST BE FALSE AND THE PREMISES TRUE.
               NOW ASSIGN CONSISTENT TRUTH VALUES TO THE REMAINING
               SENTENCE LETTERS.

T              Q
F              R
T              S

               LET US CHECK TO SEE IF YOUR TRUTH ASSIGNMENTS MAKE
               THE PREMISES TRUE AND THE CONCLUSION FALSE.

T              QVR
T              Q → S
T              R → (Q → S)
F              S → R
T              (S → R) → Q

               YOUR ASSIGNMENT MAKES ALL THE PREMISES TRUE.
               LET'S CHECK THE CONCLUSION.

T              Q&S
F              Q&S → R

CORRECT
```

505.15:

```
FIND A COUNTEREXAMPLE:      ¬ R
P              (1)          ¬ QV ¬ R
P              (2)          ¬ Q
CEX            YOUR ARE LOOKING FOR A COUNTEREXAMPLE.  THUS, THE
               CONCLUSION MUST BE FALSE AND THE PREMISES TRUE.
               NOW ASSIGN CONSISTENT TRUTH VALUES TO THE REMAINING
               SENTENCE LETTERS.

F              Q
T              R

               LET US CHECK TO SEE IF YOUR TRUTH ASSIGNMENTS MAKE
               THE PREMISES TRUE AND THE CONCLUSION FALSE.

T              ¬ Q
F              ¬ R
T              ¬ QV ¬ R

               YOUR ASSIGNMENT MAKES ALL THE PREMISES TRUE.
               LET'S CHECK THE CONCLUSION.

F              ¬ R

CORRECT
```

IV. Derive mode:

505.25:

```
DERIVE OR FIND A COUNTEREXAMPLE:   QVR
P              (1)          S → Q
P              (2)          ¬ S → R
P              (3)          S
DER

               OK...CONSTRUCT A DERIVATION OR PROOF.

1.3AA          (4)          A
4FD            (5)          (Q)V(R)

CORRECT
```

413.33:

DERIVE: $A=6 \rightarrow 5+2=A+1$

WP	(1)	$A=6$
1AE		
:1	(2)	$A+1=6+1$
2CE1	(3)	$6+1=A+1$
ND6	(4)	$6=5+1$
3.4RE1	(5)	$(5+1)+1=A+1$
5AR2	(6)	$5+(1+1)=A+1$
ND2	(7)	$2=1+1$
7CE1	(8)	$1+1=2$
6.8RE1	(9)	$5+2=A+1$
1.9CP	(10)	$A=6 \rightarrow 5+2=A+1$

CORRECT

APPENDIX B

A List of The Rules of Inference, Theorems and Axioms Used in LIS

1. Sentential Variables: Q,R,S,U,W

2. Rules of Inference:

   (a) AA: Affirm the Antecedent,
   (b) WP: Working Premise,
   (c) DN: Double Negation,
   (d) FC: Form a Conjunction,
   (e) RC: Right Conjunct,
   (f) LC: Left Conjunct,
   (g) FD: Form a Disjunct,
   (h) DD: Deny Disjunct, and
   (i) DLL: Delete last line.

3. Derivation or Proof Procedures:
   (a) CP: Conditional Proof, and
   (b) IP: Indirect Proof.

1. Numerical Variables: A,B,C,D,E.

2. Rules of Inference:
   (a) ND: Number Definition,
   (b) CE: Commute Equals,
   (c) AE: Add Equals,
   (d) SE: Subtract Equals,
   (e) LT: Rule of Logical Truth, and
   (f) RE: Replace Equals.

3. Axioms for Addition:
   - (a) CA (Commute Addition): $A+B=B+A$
   - (b) AS (Associate Addition): $(A+B)+C=A+(B+C)$
   - (c) Z (Zero Axiom): $A+0=A$
   - (d) N (Negative Number Axiom): $A+(-B)=A-B$
   - (e) AI (Additive Inverse Axiom): $A+(-A)=0$

4. Theorems on Addition:

Theorem 1: $0+A=A$

Theorem 2: $(-A)+A=0$

Theorem 3: $A-A=0$

Theorem 4: $0-A=-A$

Theorem 5: $0=-0$

Theorem 6: $A-0=A$

Theorem 7: $A+B=A+C \rightarrow B=C$

Theorem 8: $A+B=C \rightarrow A=C-B$

Theorem 9: $A=C-B \rightarrow A+B=C$

Theorem 10: $A+B=0 \rightarrow A=-B$

Theorem 11: $A=-B \rightarrow A+B=0$

Theorem 12: $A+B=A \rightarrow B=0$

Theorem 13: $-(-A)=A$

Theorem 14: $(-(A+B))+B=-A$

Theorem 15: $-(A+B)=(-A)-B$

Theorem 16: $(-A)-B=(-B)-A$

Theorem 17: $-(A-B)=B-A$

Theorem 18: $(A-B)-C=A+((-B)-C)$

Theorem 19: $(A-B)-C=A-(B+C)$

Theorem 20: $A+(B-A)=B$

Theorem 21: $A-(A+B)=-B$

Theorem 22: $(A-B)+(B-C)=A-C$

5. Additional Rules of Inference:
   - (a) ME: Multiply Equals, and
   - (b) DE: Divide Equals.

6. Axioms for Multiplication:

(a) CM (Commute Multiplication): $AXB=BXA$

(b) MS (Associate Multiplication): $(AXB)XC=AX(BXC)$

(c) MU (Multiplication by Unity): $AX1=A$

(d) MI (Multiplicative Inverse): $\neg A=0 \to AX(1/A)=1$

(e) FR (Axiom for Fraction): $\neg B=0 \to A/B=AX(1/B)$

(f) U (Unity Axiom): $\neg 1=0$

(g) DL (Distributive Law): $AX(B+C)=(AXB)+(AXC)$.

7. Theorems on Multiplication:

Theorem 30: $1XA=A$

Theorem 31: $\neg A=0 \to (1/A)XA=1$

Theorem 32: $1/1=1$

Theorem 33: $A/1=A$

Theorem 34: $\neg A=0 \to A/A=1$

Theorem 35: $\neg B=0 \& A/B=0 \to A=0XB$

Theorem 36: $(B+C)XA=(BXA)+(CXA)$

Theorem 37: $AX0=0$

Theorem 38: $\neg A=0 \to \neg 1/A=0$

Theorem 39: $\neg A=0 \to 0/A=0$

Theorem 40: $\neg A=0 \& AXB=1 \to B=1/A$

Theorem 41: $\neg A=0 \& AXB=A \to B=1$

Theorem 42: $\neg B=0 \to (A/B)XC=(AXC)/B$

Theorem 43: $\neg B=0 \to (A/B)XC=(C/B)XA$

Theorem 44: $\neg B=0 \& \neg D=0 \to (A/B)X(C/D)=(C/B)X(A/D)$

Theorem 45: $\neg A=0 \& \neg B=0 \to (A/B)X(B/A)=1$

Theorem 46: $\neg A=0 \& AXB=AXC \to B=C$

Theorem 47: $\neg A=0 \& AXB=0 \to B=0$

Theorem 48: $\neg AXB=0 \to \neg A=0 \& \neg B\ 0$

Theorem 49: $\neg A=0 \& \neg B=0 \to \neg AXB=0$

Theorem 50: $\neg A=0 \& \neg B=0 \to B/(AXB)=1/A$

Theorem 51: $\neg A=0 \& \neg B=0 \to (CXB)/(AXB)=C/A$

Theorem 52: $(\neg B=0 \& \neg D=0) \& A/B=C/D \to AXD=CXB$

Theorem 53: $\neg B=0 \& A=BXC \to A/B=C$

Theorem 54: $AX(-B)=-(AXB)$

Theorem 55: $(-A)X(-B)=AXB$

8. Ordering Axioms:

(a) NS (Asymmetry): $A < B \to \neg B < A$

(b) AD (additivity): $A < B \to A+C < B+C$

(c) MD (Multiplicativity): $A < B \& 0 < C \to AXC < BXC$

(d) TR (transitivity): $A < B \& B < C \to A < C$

(e) CN (connectivity): $A \neq B \to A < B V B < A$

9. Theorems on Inequalities:

Theorem 60: $\neg A < A$

Theorem 61: $A=B \to \neg A < B \& \neg B < A$

Theorem 62: $A < B \to \neg A=B \& \neg B < A$

Theorem 63: $A < 0 \to 0 < -A$

Theorem 64: $0 < -A \to A < 0$

Theorem 65: $A+B < A+C \to B < C$

Theorem 66: $A < B \to -B < -A$

Theorem 67: $-B < -A \to A < B$

Theorem 68: $A+(-B) < A+(-C) \to C < B$

Theorem 69: $C < B \to A+(-B) < A+(-C)$

Theorem 70: $A < 0 \& B < C \to AXC < AXB$

Theorem 71: $A < 0 \& AXB < AXC \to C < B$

Theorem 72: $0 < A \& AXB < AXC \to B < C$

Theorem 73 $0 < 1$

Theorem 74: $A < 0 \to 1/A < 0$

Theorem 75: $0 < A \& (B < 0 \& C < 0) \to AXB < BXC$

Theorem 76: $A < 0 \& (0 < B \& 0 < C) \to AXB < BXC$

Theorem 77: $\neg B=0 \& 0 < A/B \to 0 < AXB$

Theorem 78: $\neg B=0 \& 0 < AXB \to 0 < A/B$

Boolean or Class Algebra

1. Class Variables: G, H, M, K, L

2. Axioms:

(a) CU (Commute Union): $G \cup H = H \cup G$

(b) CI (Commute Intersection): $G \cap H = H \cap G$

(c) UI (Union Identity): $G \cup 0 = G$

(d) II (Intersection Identity): $G \cap X = G$

(e) DU (Distribute Union): $G \cup (H \cap K) = (G \cup H) \cap (G \cup K)$

(f) DI (Distribute Intersection): $G \cap (H \cup K) = (G \cap H) \cup (G \cap K)$

(g) EM (Excluded Middle): $G \cup (-G) = X$

(h) RD (Reduction): $G \cap (-G) = 0$

(i) UC (Associate Union): $(G \cup H) \cup K = G \cup (H \cup K)$

(j) IA (Associate Intersection): $(G \cap H) \cap K = G \cap (H \cap K)$

(k) SA (Subclass Axiom): $G \cap (-H) = 0 \rightarrow G \subset H$

(l) CS (Converse of Subclass): $G \subset H \rightarrow G \cap (-H) = 0$

3. Theorems:

Theorem 161: $G \cup ((-G) \cap H) = G \cup H$

Theorem 162: $G \cap ((-G) \cup H) = G \cap H$

Theorem 163: $G \cup G = G$

Theorem 164: $G \cap G = G$

Theorem 165: $G \cup X = X$

Theorem 166: $G \cap 0 = 0$

Theorem 167: $G \cup (G \cap H) = G$

Theorem 168: $G \cap (G \cup H) = G$

Theorem 169: $G \cap (-H) = 0 \& G \cap H = 0 \rightarrow G = 0$

Theorem 170: $G \cup (-H) = X \& G \cup H = X \rightarrow G = X$

Theorem 171: $G \cup H = 0 \rightarrow G = 0$

Theorem 172: $G \cap H = X \rightarrow G = X$

Theorem 173: $G \cup H = G \cup K \& G \cap H = G \cap K \rightarrow H = K$

Theorem 174: $(G \cup H = X \& G \cup K = X) \& (G \cap H = 0 \& G \cap K = 0) \rightarrow H = K$

Theorem 175: $(G \cup H = G \& G \cup K = G) \& (G \cap H = 0 \& G \cap K = 0) \rightarrow H = K$

Theorem 176: $(G \cup H = X \& G \cup K = X) \& (G \cap H = G \& G \cap K = G) \rightarrow H = K$

Theorem 177: $-(-G) = G$

Theorem 178: $-X=0$

Theorem 179: $G \cup H = G \cap H \rightarrow G = H$

Theorem 180: $G \cap (H \cap K) = (G \cap H) \cap (G \cap K)$

Theorem 190: $G \subset G$

Theorem 191: $0 \subset G$

Theorem 192: $G \subset X$

Theorem 193: $G \subset H \& H \subset G \rightarrow G = H$

Theorem 194: $G \subset H \rightarrow G \cup H = H$

Theorem 195: $G \cup H = H \rightarrow G \subset H$

Theorem 196: $G \cup (-H) = X \rightarrow G \subset H$

Theorem 197: $G \subset H \rightarrow G \cup (-H) = X$

Theorem 198: $G \subset H \rightarrow G \cap H = G$

Theorem 199: $G \cap H = G \rightarrow G \subset H$

Theorem 200: $G \subset H \& H \subset K \rightarrow G \subset K$

Theorem 201: $G \subset H \rightarrow -H \subset -G$

Theorem 202: $G \subset H \& G \subset -H \rightarrow G = 0$

Theorem 203: $G \subset H \& -G \subset H \rightarrow H = X$

Theorem 204: $G \subset G \cup H$

Theorem 205: $G \cap H \subset G$

Theorem 206: $G \subset K \& H \subset K \rightarrow G \cup H \subset K$

Theorem 207: $G \subset H \& G \subset K \rightarrow G \subset H \cap K$

Theorem 208: $G \subset H \rightarrow H = G \cup (H \cap (-G))$

APPENDIX C

Two Examples of Derivation Problems from LIS

This appendix contains two examples of derivation problems from LIS. Example 1 is typical of the sentential logic problems. Example 2 is typical of the algebra problems.

An explanation of the lines of the derivation in Example 1 follows:

(1) - (5) These are the given premises to be used in deriving the logical sentence R.

(6) The student introduces the denial of the sentence to be derived. To do this, he uses the working premise rule, WP. LIS indents this premise and all lines following it until the student proves a contradiction and uses the indirect proof rule, IP, to derive the denial of what he entered on this line. See the explanation for line 14 (below).

(7) Line 1 is a disjunction and the newly introduced line 6 is the denial of one of the disjuncts. The DD rule (Deny Disjunct) allows the student to establish the truth of the other disjunct S.

(8) Line 2 is the conditional "if not Q, then not S." Line 7 states that S is true, so the student used deny consequent, DC, to prove that Q is true.

(9) The antecedent of the conditional in the line 3 premise is in the form of a double negation (not (not Q)): the student has proved that Q is true in line 8, so he uses double negation, DN, to derive this antecedent.

(10) Now he uses the affirm the antecedent rule, AA, to derive the consequent of line 3.

(11) He uses double negation again, now on the premise line 4.

(12) He uses affirm the antecedent again to derive not W.

(13) He uses deny disjunct again, this time on the disjunct on line 12 to get not S.

(14) He has derived a contradiction with the help of the working premises he introduced. On line 7 he has S is true. On line 13 he has not S is true. He uses the indirect proof rule, IP, to establish the denial of not R, the working premise on line 6.

Insert Table 1 about here

Now we give a detailed explanation of the steps in the derivation of Example 2. There are no premises and the student is being asked to prove Theorem 22 which will then become available to him for use in later proofs.

(1) The student introduces the negative number axiom, N. The computer prints out the axiom and then allows the student to substitute expressions for A and B. In this case, the student types A for A and B for B.

(2) Line 1 is an equation, so the student can commute about =. To do this, he uses the commute equals rule, CE, where the first 1 is the line number and the second 1 is the occurrence number of the =.

(3) The student wishes to add something to both sides of the equality on line 2. To do this, he uses the add equals rule, AE, where the 2 is the line number of the equation. The computer types a colon after which the students types the expression to be added. The computer then types line 3.

On the next line the student attempts to type a rule which the computer does not recognize.

(4) The student again uses the negative number axiom.

(5) He applies CE to line 4.

(6) The student now uses the replace equals rule, RE. He wishes to replace an occurrence in line 3 of the left-hand side of the equation in line 5 by the right-hand side of the equation in line 5. There is more than one occurrence of (B-C) in line 3 and the student specifies which one he wants replaced by the occurrence number, 1.

On the next line he decides to erase line 6. He does this by using DLL, delete the last line.

(7) He again uses RE, this time for the second occurrence of (B-C). The student wishes to associate addition to the right in line 6. To do this, he uses the associate right rule, AR. He wants to associate about the second plus sign, hence he uses 2 as the occurrence number. Since this is not possible, he receives an error message.

(8) He again tries AR, only this time the occurrence number of the plus sign is 3.

(9) He associates left about the third plus sign using AL.

(10) He uses the negative number axiom again.

(11) He now makes use of a theorem which he had proved earlier. A theorem is used in a manner analogous to the axioms. On the next line he misuses RE and receives the appropriate error message.

(12) He correctly uses RE.

(13) He makes use of theorem 1.

(14)-(15) Two more uses of RE establish the desired theorem.

Insert Table 2 about here

TABLE 1

Example 1

```
400.1

DERIVE R

P            (1)        RVS
P            (2)        ¬ Q →  ¬ S
P            (3)        ¬ (¬ Q) → (U →¬ W)
P            (4)        ¬ (¬ U)
P            (5)        WV ¬ S
WP           (6)                               ¬ R
1.6DD        (7)                               S
2.7DC        (8)                               Q
8DN          (9)                               ¬ (¬ Q)
3.9AA       (10)                               U → ¬ W
4DN         (11)                               U
10.11AA     (12)                               ¬ W
5.12DD      (13)                               ¬ S
6.7.13IP    (14)        R
CORRECT
```

TABLE 2

Example 2

417.31:

HERE IS THEOREM 22
DERIVE: (A-B) + (B-C) = A-C

N A + (-B) = A-B		
A:A		
B:B	(1)	A + (-B) = A-B
1CE1	(2)	A-B = A + (-B)
2AE		
: B-C	(3)	(A-B) + (B-C) = (A + (-B)) + (B-C)
NU		
	NU IS NOT A RULE	
N A + (-B) = A-B		
A:B		
B:C	(4)	B + (-C) = B-C
4CE1	(5)	B-C = B + (-C)
3.5RE1	(6)	(A-B) + (B + (-C)) = (A + (-B)) + (B-C)
DLL		
3.5RE2	(6)	(A-B) + (B-C) = (A +(-B)) + (B + (-C))
6AR2		
	YOU MAY NOT ASSOCIATE RIGHT AROUND + SIGN NUMBER 2	
6AR3	(7)	(A-B) + (B-C) = A + ((-B) + (B + (-C)))
7AL3	(8)	(A-B) + (B-C) = A + (((-B) + B) + (-C))
N A + (-B) = A-B		
A: A		
B: C	(9)	A + (-C) = A-C
TH2 (-A) + A = 0		
A:B	(10)	(-B) + B = 0
8.10RE2		
	THERE ARE NOT 2 OCCURRENCES OF (-B) + B IN LINE 8	

```
8.10RE1
               (11)        (A-B) + (B-C) = A + (0 + (-C))
TH2  0 + A = A
A: -C          (12)        0 + (-C) = -C
11.12RE1
               (13)        (A-B) + (B-C) = A + (-C)
13.9RE1        (14)        (A-B) + (B-C) = A-C
CORRECT
```

APPENDIX D

Description of Data Analysis Programs

In this appendix we describe the programs, written by the author, which were used to reduce and analyze the data.

Logic Program

Each day during the summer of 1970, a file was created for each logic student on the PDP-10 disk file system. Files were identified on the disk by a file name (up to six characters) and a file extension (up to three characters) written as NNNNNN.EEE. The name chosen for each student file was the student's account number, the extension was the date. Thus, logic student L1125 on July 13 had his data recorded on a file named L1125.713. At the end of each day, the student data files were transferred to magnetic tape. The format of these files is given in Table 1.

Insert Table 1 about here

Data Reduction

In the fall of 1970, a series of programs were written to convert the raw data into a format acceptable to the BMD programs. We give here a brief description of these programs, indicating the programming language used in each case.

PASS1 - PDP-10 assembly language

Input: daily student data files

(1) combined the data in the individual student files described above into one data file per student.

PASS2 - SAIL*

Input: output files from PASS1

(1) created a separate file for each logic problem.

PASS3 - SAIL*

Input: output files from PASS2

(1) Extracted the following information from each problem file:

(a) problem number

(b) number of students who attempted the problem

(c) number of students for whom there was complete data on the problem. As mentioned in Chapter III, some data were lost due to system or machine failures, so that there were incomplete data for some students on some problems.

(d) mean and standard deviation of the number of lines in a complete derivation for the problem. Here and below we define the mean as:

$$\text{Mean} = \bar{X} = \left(\sum_{i=1}^{N} X_i\right)/N$$

and the standard deviation as:

$$\text{Stan. Dev.} = \sqrt{\left(\sum_{i=1}^{N} (X_i - \bar{X})^2\right)/N-1}$$

where N is the number of students completing the problem.

(e) mean and standard deviation of latency to solution.

(f) mean and standard deviation of latency per line.

(g) mean and standard deviation of corrected latency per line.

---

*Stanford Artificial Intelligence Laboratory's Algol-like language.

(h) mean and standard deviation of number of error messages.

(i) mean and standard deviation of number of DLL's.

(j) mean and standard deviation of number of restarts.

(2) created ASCII files of the above information formated for printing on a teletype or displaying on a CRT. These could also be used as input for the BMD programs.

COMB - Fortran

Input: output from PASS3 and a file containing the values of the structural variables which were typed as input by hand on the CRT's.

(1) combined the two input files into one file containing both the behavioral and structural variables.

SORT - Fortran

Input: output from PASS3

(1) produced a rank-ordering of the problems for each of the five behavioral measures.

Analysis

In addition to writing the above programs, I also implemented the BMD06M program on the PDP-10 and modified the already existing BMD02R program to produce the plots mentioned in Chapter III.

TABLE 1

Format of Raw Logic Data

The first four words of each student file were:

word1: Student account number
word2: Date
word3: Start time
word4: New day code - 761616161616

Whenever the student was restarted, the above four words were put in his file.

The first words for every problem were:

word1: New problem code - 716161616161
word2: Problem start time
word3: Problem and lesson number
words 4-n: Problem type codes

These were followed by response codes. For each student input these were:

word1: response code - 767676767676
words 2-n-1: Student response in ASCII
wordn: Latency to response

Each time a student timed out, the following information was recorded:

word1: TIMOUT
word2: Time of the time-out

Each time a student asked for a hint and the hint clock had not fired, the student received one of the following two messages. For "A HINT IS NOT AVAILABLE NOW" we recorded:

word1: NOTNOW
word2: Time of message

When the student received "THINK A LITTLE LONGER", we recorded:

word1: KEEPON
word2: Time of message

When the student received an error message, we recorded:

word1: ERRORS
word2: Error message number
word3: Contents of an accumulator containing information about the error
word4: Time of the error

At the end of each problem, we recorded:

word1: Problem end code- 766766766766
word2: Time of end of problem

Finally, at the time that each student was signed off, we recorded:

word1: Sign-off code - 776776776776
word2: Time of sign-off
word3: 747474747474

(Continued from inside front cover)

96 R. C. Atkinson, J. W. Brelsford, and R. M. Shiffrin. Multi-process models for memory with applications to a continuous presentation task. April 13, 1966. (*J. math. Psychol.*, 1967, *4*, 277-300).
97 P. Suppes and E. Crothers. Some remarks on stimulus-response theories of language learning. June 12, 1966.
98 R. Bjork. All-or-none subprocesses in the learning of complex sequences. (*J. math. Psychol.*, 1968, *1*, 182-195).
99 E. Gammon. The statistical determination of linguistic units. July 1, 1966.
100 P. Suppes, L. Hyman, and M. Jerman. Linear structural models for response and latency performance in arithmetic. (In J. P. Hill (ed.), *Minnesota Symposia on Child Psychology*. Minneapolis, Minn.: 1967. Pp. 160-200).
101 J. L. Young. Effects of intervals between reinforcements and test trials in paired-associate learning. August 1, 1966.
102 H. A. Wilson. An investigation of linguistic unit size in memory processes. August 3, 1966.
103 J. T. Townsend. Choice behavior in a cued-recognition task. August 8, 1966.
104 W. H. Batchelder. A mathematical analysis of multi-level verbal learning. August 9, 1966.
105 H. A. Taylor. The observing response in a cued psychophysical task. August 10, 1966.
106 R. A. Bjork. Learning and short-term retention of paired associates in relation to specific sequences of interpresentation intervals. August 11, 1966.
107 R. C. Atkinson and R. M. Shiffrin. Some Two-process models for memory. September 30, 1966.
108 P. Suppes and C. Ihrke. Accelerated program in elementary-school mathematics--the third year. January 30, 1967.
109 P. Suppes and I. Rosenthal-Hill. Concept formation by kindergarten children in a card-sorting task. February 27, 1967.
110 R. C. Atkinson and R. M. Shiffrin. Human memory: a proposed system and its control processes. March 21, 1967.
111 Theodore S. Rodgers. Linguistic considerations in the design of the Stanford computer-based curriculum in initial reading. June 1, 1967.
112 Jack M. Knutson. Spelling drills using a computer-assisted instructional system. June 30, 1967.
113 R. C. Atkinson. Instruction in initial reading under computer control: the Stanford Project. July 14, 1967.
114 J. W. Brelsford, Jr. and R. C. Atkinson. Recall of paired-associates as a function of overt and covert rehearsal procedures. July 21, 1967.
115 J. H. Stelzer. Some results concerning subjective probability structures with semiorders. August 1, 1967
116 D. E. Rumelhart. The effects of interpresentation intervals on performance in a continuous paired-associate task. August 11, 1967.
117 E. J. Fishman, L. Keller, and R. E. Atkinson. Massed vs. distributed practice in computerized spelling drills. August 18, 1967.
118 G. J. Green. An investigation of some counting algorithms for simple addition problems. August 21, 1967.
119 H. A. Wilson and R. C. Atkinson. Computer-based instruction in initial reading: a progress report on the Stanford Project. August 25, 1967.
120 F. S. Roberts and P. Suppes. Some problems in the geometry of visual perception. August 31, 1967. (*Synthese*, 1967, *17*, 173-201)
121 D. Jamison. Bayesian decisions under total and partial ignorance. D. Jamison and J. Kozielecki. Subjective probabilities under total uncertainty. September 4, 1967.
122 R. C. Atkinson. Computerized instruction and the learning process. September 15, 1967.
123 W. K. Estes. Outline of a theory of punishment. October 1, 1967.
124 T. S. Rodgers. Measuring vocabulary difficulty: An analysis of item variables in learning Russian-English and Japanese-English vocabulary parts. December 18, 1967.
125 W. K. Estes. Reinforcement in human learning. December 20, 1967.
126 G. L. Wolford, D. L. Wessel, W. K. Estes. Further evidence concerning scanning and sampling assumptions of visual detection models. January 31, 1968.
127 R. C. Atkinson and R. M. Shiffrin. Some speculations on storage and retrieval processes in long-term memory. February 2, 1968.
128 John Holmgren. Visual detection with imperfect recognition. March 29, 1968.
129 Lucille B. Mlodnosky. The Frostig and the Bender Gestalt as predictors of reading achievement. April 12, 1968.
130 P. Suppes. Some theoretical models for mathematics learning. April 15, 1968. (*Journal of Research and Development in Education*, 1967, *1*, 5-22)
131 G. M. Olson. Learning and retention in a continuous recognition task. May 15, 1968.
132 Ruth Norene Hartley. An investigation of list types and cues to facilitate initial reading vocabulary acquisition. May 29, 1968.
133 P. Suppes. Stimulus-response theory of finite automata. June 19, 1968.
134 N. Moler and P. Suppes. Quantifier-free axioms for constructive plane geometry. June 20, 1968. (In J. C. H. Gerretsen and F. Oort (Eds.), *Compositio Mathematica*. Vol. 20. Groningen, The Netherlands: Wolters-Noordhoff, 1968. Pp. 143-152.)
135 W. K. Estes and D. P. Horst. Latency as a function of number of response alternatives in paired-associate learning. July 1, 1968.
136 M. Schlag-Rey and P. Suppes. High-order dimensions in concept identification. July 2, 1968. (*Psychon. Sci.*, 1968, *11*, 141-142)
137 R. M. Shiffrin. Search and retrieval processes in long-term memory. August 15, 1968.
138 R. D. Freund, G. R. Loftus, and R.C. Atkinson. Applications of multiprocess models for memory to continuous recognition tasks. December 18, 1968.
139 R. C. Atkinson. Information delay in human learning. December 18, 1968.
140 R. C. Atkinson, J. E. Holmgren, and J. F. Juola. Processing time as influenced by the number of elements in the visual display. March 14, 1969.
141 P. Suppes, E. F. Loftus, and M. Jerman. Problem-solving on a computer-based teletype. March 25, 1969.
142 P. Suppes and Mona Morningstar. Evaluation of three computer-assisted instruction programs. May 2, 1969.
143 P. Suppes. On the problems of using mathematics in the development of the social sciences. May 12, 1969.
144 Z. Domotor. Probabilistic relational structures and their applications. May 14, 1969.
145 R. C. Atkinson and T. D. Wickens. Human memory and the concept of reinforcement. May 20, 1969.
146 R. J. Titiev. Some model-theoretic results in measurement theory. May 22, 1969.
147 P. Suppes. Measurement: Problems of theory and application. June 12, 1969.
148 P. Suppes and C. Ihrke. Accelerated program in elementary-school mathematics--the fourth year. August 7, 1969.
149 D. Rundus and R.C. Atkinson. Rehearsal in free recall: A procedure for direct observation. August 12, 1969.
150 P. Suppes and S. Feldman. Young children's comprehension of logical connectives. October 15, 1969.

(Continued on back cover)

(Continued from inside back cover)

151 Joaquim H. Laubsch. An adaptive teaching system for optimal item allocation. November 14, 1969.

152 Roberta L. Klatzky and Richard C. Atkinson. Memory scans based on alternative test stimulus representations. November 25, 1969.

153 John E. Holmgren. Response latency as an indicant of information processing in visual search tasks. March 16, 1970.

154 Patrick Suppes. Probabilistic grammars for natural languages. May 15, 1970.

155 E. Gammon. A syntactical analysis of some first-grade readers. June 22, 1970.

156 Kenneth N. Wexler. An automaton analysis of the learning of a nimiature system of Japanese. July 24, 1970.

157 R. C. Atkinson and J.A. Paulson. An approach to the psychology of instruction. August 14, 1970.

158 R.C. Atkinson, J.D. Fletcher, H.C. Chetin, and C.M. Stauffer. Instruction in initial reading under computer control: the Stanford project. August 13, 1970.

159 Dewey J. Rundus. An analysis of rehearsal processes in free recall. August 21, 1970.

160 R.L. Klatzky, J.F. Juola, and R.C. Atkinson. Test stimulus representation and experimental context effects in memory scanning.

161 William A. Rottmayer. A formal theory of perception. November 13, 1970.

162 Elizabeth Jane Fishman Loftus. An analysis of the structural variables that determine problem-solving difficulty on a computer-based teletype. December 18, 1970.

163 Joseph A. Van Campen. Towards the automatic generation of programmed foreign-language instructional materials. January 11, 1971.

164 Jamesine Friend and R.C. Atkinson. Computer-assisted instruction in programming: AID. January 25, 1971.

165 Lawrence James Hubert. A formal model for the perceptual processing of geometric configurations. February 19, 1971.

166 J. F. Juola, I.S. Fischler, C.T. Wood, and R.C. Atkinson. Recognition time for information stored in long-term memory.

167 R.L. Klatzky and R.C. Atkinson. Specialization of the cerebral hemispheres in scanning for information in short-term memory.

168 J.D. Fletcher and R.C. Atkinson. An evaluation of the Stanford CAI program in initial reading (grades K through 3). March 12, 1971.

169 James F. Juola and R.C. Atkinson. Memory scanning for words versus categories.

170 Ira S. Fischler and James F. Juola. Effects of repeated tests on recognition time for information in long-term memory.

171 Patrick Suppes. Semantics of context-free fragments of natural languages. March 30, 1971.

172 Jamesine Friend. Instruct coders' manual. May 1, 1971.

173 R.C. Atkinson and R.M. Shiffrin. The control processes of short-term memory. April 19, 1971.

174 Patrick Suppes. Computer-assisted instruction at Stanford. May 19, 1971.

175 D. Jamison, J.D. Fletcher, P. Suppes and R.C.Atkinson. Cost and performance of computer-assisted instruction for compensatory education.

176 Joseph Offir. Some mathematical models of individual differences in learning and performance. June 28, 1971.

177 Richard C. Atkinson and James F. Juola. Factors influencing speed and accuracy of word recognition. August 12, 1971.

178 P. Suppes, A. Goldberg, G. Kanz, B. Searle and C. Stauffer. Teacher's handbook for CAI courses. September 1, 1971.

179 Adele Goldberg. A generalized instructional system for elementary mathematical logic. October 11, 1971.

180 Max Jerman. Instruction in problem solving and an analysis of structural variables that contribute to problem-solving difficulty. November 12, 1971

181 Patrick Suppes. On the grammar and model-theoretic semantics of children's noun phrases. November 29, 1971.

182 Georg Kreisel. Five notes on the application of proof theory to computer science. December 10, 1971.

183 James Michael Moloney. An investigation of college student performance on a logic curriculum in a computer-assisted instruction setting. January 28, 1972.